Fig. 1A

```
                       10        20        30        40        50        60
                        *         *         *         *         *         *
BTHKURHD      ATGGATAACAATCCGAACATCAATGAATGCATTCCTTATAATTGTTTAAGTAACCCTGAA
flsynbt.fin   .....C.....C..C........C..G.....C..C..C..C..CC.G..C.....C..G
bssyn         .....C.....C..C........C..G.....C..C..C..C..CC.G..C.....C..G

                       70        80        90       100       110       120
                        *         *         *         *         *         *
BTHKURHD      GTAGAAGTATTAGGTGGAGAAAGAATAGAAACTGGTTACACCCCAATCGATATTTCCTTG
flsynbt.fin   ..G..G..GC.G..C..C..GC.C..C..G..C..C........C.....C..CAG.C..
bssyn         ..G..G..GC.G..C..C..GC.C..C..G..C..C........C.....C..CAG.C..

                      130       140       150       160       170       180
                        *         *         *         *         *         *
BTHKURHD      TCGCTAACGCAATTTCTTTTGAGTGAATTTGTTCCCGGTGCTGGATTTGTGTTAGGACTA
flsynbt.fin   AGC..G..C..G..C..GC....C..G..C..G.....C..C..C..C...C.G..C..G
bssyn         AGC..G..C..G..C..GC....C..G..C..G.....C..C..C..C...C.G..C..G

                      190       200       210       220       230       240
                        *         *         *         *         *         *
BTHKURHD      GTTGATATAATATGGGGAATTTTTGGTCCCTCTCAATGGGACGCATTTCTTGTACAAATT
flsynbt.fin   ..G..C..C..C.....C..C..C..C...AGC..G........C..C..G..G..G..C
bssyn         ..G..C..C..C.....C..C..C..C...AGC..G........C..C..G..G..G..C

                      250       260       270       280       290       300
                        *         *         *         *         *         *
BTHKURHD      GAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAGGAACCAAGCCATTTCTAGATTA
flsynbt.fin   ..G...C.G..C.....GC.C..C..G..G.....CC.C.....G.....CAGCC.CC.G
bssyn         ..G...C.G..C.....GC.C..C..G..G.....CC.C.....G.....CAGCC.CC.G

                      310       320       330       340       350       360
                        *         *         *         *         *         *
BTHKURHD      GAAGGACTAAGCAATCTTTATCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAGAT
flsynbt.fin   ..G..C..G.....C..G..C.....C.....C..GAGC..CC.C........G..C..C
bssyn         ..G..C..G.....C..G..C.....C.....C..GAGC..CC.C........G..C..C

                      370       380       390       400       410       420
                        *         *         *         *         *         *
BTHKURHD      CCTACTAATCCAGCATTAAGAGAAGAGATGCGTATTCAATTCAATGACATGAACAGTGCC
flsynbt.fin   ..C..C..C..C..CC.GC.C..G........C..C..G.....C...........C...
bssyn         ..C..C..C..C..CC.GC.C..G........C..C..G.....C...........C...

                      430       440       450       460       470       480
                        *         *         *         *         *         *
BTHKURHD      CTTACAACCGCTATTCCTCTTTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTA
flsynbt.fin   ..G..C.....C..C..C..G..C..C..G..G..C..C..G..G..C..GC.GAGC..G
bssyn         ..G..C.....C..C..C..G..C..C..G..G..C..C..G..G..C..GC.GAGC..G

                      490       500       510       520       530       540
                        *         *         *         *         *         *
BTHKURHD      TATGTTCAAGCTGCAAATTTACATTTATCAGTTTTGAGAGATGTTTCAGTGTTTGGACAA
flsynbt.fin   ..C..G..G..C..C..CC.G..CC.GAGC..GC..C.C..C..CAGC.....C..C..G
bssyn         ..C..G..G..C..C..CC.G..CC.GAGC..GC..C.C..C..CAGC.....C..C..G

                      550       560       570       580       590       600
                        *         *         *         *         *         *
BTHKURHD      AGGTGGGGATTTGATGCCGCGACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT
flsynbt.fin   C.C.....C..C..C.....C..C.....C..C..C..C..C..CC.G..CC.C..G..C
bssyn         C.C.....C..C..C.....C..C.....C..C..C..C..C..CC.G..CC.C..G..C
```

Fig. 1B

```
                         610       620       630       640       650       660
                           *         *         *         *         *         *
BTHKURHD        GGCAACTATACAGATCATGCTGTACGCTGGTACAATACGGGATTAGAGCGTGTATGGGGA
flsynbt.fin     ........C..C..C..C..C..G...........C..C..CC.G.....C..G.....T
bssyn           ........C..C..C..C..C..G...........C..C..CC.G.....C..G.....T

                         670       680       690       700       710       720
                           *         *         *         *         *         *
BTHKURHD        CCGGATTCTAGAGATTGGATAAGATATAATCAATTTAGAAGAGAATTAACACTAACTGTA
flsynbt.fin     ..C..CAGCC.C..C.....C..G..C..C..G..CC.CC.C..GC.G..C..G..C..G
bssyn           ..C..CAGCC.C..C.....C..G..C..C..G..CC.CC.C..GC.G..C..G..C..G

                         730       740       750       760       770       780
                           *         *         *         *         *         *
BTHKURHD        TTAGATATCGTTTCTCTATTTCCGAACTATGATAGTAGAACGTATCCAATTCGAACAGTT
flsynbt.fin     C.G..C.....GAGC..G..C..C.....C..C..CC.C..C..C..C..C..C..C..G
bssyn           C.G..C.....GAGC..G..C..C.....C..C..CC.C..C..C..C..C..C..C..G

                         790       800       810       820       830       840
                           *         *         *         *         *         *
BTHKURHD        TCCCAATTAACAAGAGAAATTTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTTTT
flsynbt.fin     AG...GC.G..CC.C..G.....C..C.....C..GC.G..G..C..C..C..C..C..C
bssyn           AG...GC.G..CC.C..G.....C..C.....C..GC.G..G..C..C..C..C..C..C

                         850       860       870       880       890       900
                           *         *         *         *         *         *
BTHKURHD        CGAGGCTCGGCTCAGGGCATAGAAGGAAGTATTAGGAGTCCACATTTGATGGATATACTT
flsynbt.fin     ..C...AGC..C........C..G..C..C..CC.C..C..C..CC.......C..C..G
bssyn           ..C...AGC..C........C..G..C..C..CC.C..C..C..CC.......C..C..G

                         910       920       930       940       950       960
                           *         *         *         *         *         *
BTHKURHD        AACAGTATAACCATCTATACGGATGCTCATAGAGGAGAATATTATTGGTCAGGGCATCAA
flsynbt.fin     .....C..C........C..C..C..C..CC.C..C..G..C..C...AGC..C..C..G
bssyn           .....C..C........C..C..C..C..CC.C..C..G..C..C...AGC..C..C..G

                         970       980       990      1000      1010      1020
                           *         *         *         *         *         *
BTHKURHD        ATAATGGCTTCTCCTGTAGGGTTTTCGGGGCCAGAATTCACTTTTCCGCTATATGGAACT
flsynbt.fin     ..C.....CAGC..C..C..C..CAGC..C..C..G.....C..C..C..G..C..C..C
bssyn           ..C.....CAGC..C..C..C..CAGC..C..C..G.....C..C..C..G..C..C..C

                        1030      1040      1050      1060      1070      1080
                           *         *         *         *         *         *
BTHKURHD        ATGGGAAATGCAGCTCCACAACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATAGA
flsynbt.fin     .....C..C..T..A..T..G..G..C..C..G..A..G..G..C.....A.....CC.C
bssyn           .....C..C..T..A..T..G..G..C..C..G..A..G..G..C.....A.....CC.C

                        1090      1100      1110      1120      1130      1140
                           *         *         *         *         *         *
BTHKURHD        ACATTATCGTCCACTTTATATAGAAGACCTTTTAATATAGGGATAAATAATCAACAACTA
flsynbt.fin     ..CC.GAGCAG...CC.G..CC.TC.......C..C..C..C..C..C..C..G..G..G
bssyn           ..CC.GAGCAG...CC.G..CC.TC.......C..C..C..C..C..C..C..G..G..G

                        1150      1160      1170      1180      1190      1200
                           *         *         *         *         *         *
BTHKURHD        TCTGTTCTTGACGGGACAGAATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTA
flsynbt.fin     AGC..G..G.....C..C..G..C..C..C..C...AG.AGC..CC....CAG...C..G
bssyn           AGC..G..G.....C..C..G..C..C..C..C...AG.AGC..CC....CAG...C..G
```

Fig. 1C

```
                          1210      1220      1230      1240      1250      1260
                             *         *         *         *         *         *
BTHKURHD      TACAGAAAAAGCGGAACGGTAGATTCGCTGGATGAAATACCGCCACAGAATAACAACGTG
flsynbt.fin   ...C.C..G.....C..C..G..CAGC.....C..G..C..C..T.....C.........
bssyn         ...C.C..G.....C..C..G..CAGC.....C..G..C..C..T.....C.........

                          1270      1280      1290      1300      1310      1320
                             *         *         *         *         *         *
BTHKURHD      CCACCTAGGCAAGGATTTAGTCATCGATTAAGCCATGTTTCAATGTTTCGTTCAGGCTTT
flsynbt.fin   ......C.A..G..C..C..C..C..TC.G.....C..GAGC.....C..CAGT.....C
bssyn         ......C.A..G..C..C..C..C..TC.G.....C..GAGC.....C..CAGT.....C

                          1330      1340      1350      1360      1370      1380
                             *         *         *         *         *         *
BTHKURHD      AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCTCTTGGATACATCGTAGTGCT
flsynbt.fin   ..C..C..C..C..G..C..C..CC.T..A.........AGC.....T..C..C.....C
bssyn         ..C..C..C..C..G..C..C..CC.T..A.........AGC.....T..C..C.....C

                          1390      1400        1410      1420      1430
                             *         *           *         *         *
BTHKURHD      GAATTTAATAATATAATTCCTTCATCA--CAAATTACACAAATACCTTTAACAAAATCTA
flsynbt.fin   ..G..C..C..C..C..C..--..G..GC..G..C..C..G..C..CC.G..C..GAGC.
bssyn         ..G..C..C..C..C..C..--..G..GC..G..C..C..G..C..CC.G..C..GAGC.

           1440      1450      1460      1470      1480      1490
              *         *         *         *         *         *
BTHKURHD      CTAATCTTGGCTCTGGAACTTCTGTCGTTAAAGGACCAGGATTTACAGGAGGAGATATTC
flsynbt.fin   .C..C..G...AGC..C..CAGC..G..G..G..C..C..C..C..C..C..C..C.
bssyn         .C..C..G...AGC..C..CAGC..G..G..G..C..C..C..C..C..C..C..C.

           1500      1510      1520      1530      1540      1550
              *         *         *         *         *         *
BTHKURHD      TTCGAAGAACTTCACCTGGCCAGATTTCAACCTTAAGAGTAAATATTACTGCACCATTAT
flsynbt.fin   .G..CC.C..CAGC..C........CAGC...C.GC.C..G..C..C..C..C..CC.GA
bssyn         .G..CC.C..CAGC..C........CAGC...C.GC.C..G..C..C..C..C..CC.GA

           1560      1570      1580      1590      1600      1610
              *         *         *         *         *         *
BTHKURHD      CACAAAGATATCGGGTAAGAATTCGCTACGCTTCTACCACAAATTTACAATTCCATACAT
flsynbt.fin   GC..GC.C..C..C..CC.C..C........CAGC.....C..CC.G..G.....C..CA
bssyn         GC..GC.C..C..C..CC.C..C........CAGC.....C..CC.G..G.....C..CA

           1620      1630      1640      1650      1660      1670
              *         *         *         *         *         *
BTHKURHD      CAATTGACGGAAGACCTATTAATCAGGGGAATTTTTCAGCAACTATGAGTAGTGGGAGTA
flsynbt.fin   GC..C.....CC.C..C..C..C.....C..C..CAGC..C..C.....C..C..C..C.
bssyn         GC..C.....CC.C..C..C..C.....C..C..CAGC..C..C.....C..C..C..C.

           1680      1690      1700      1710      1720      1730
              *         *         *         *         *         *
BTHKURHD      ATTTACAGTCCGGAAGCTTTAGGACTGTAGGTTTTACTACTCCGTTTAACTTTTCAAATG
flsynbt.fin   .CC.G...AG...C.....CC.C..C..G..C..C..C..C..C..C.....CAGC..C.
bssyn         .CC.G...AG...C.....CC.C..C..G..C..C..C..C..C..C.....CAGC..C.

           1740      1750      1760      1770      1780      1790
              *         *         *         *         *         *
BTHKURHD      GATCAAGTGTATTTACGTTAAGTGCTCATGTCTTCAATTCAGGCAATGAAGTTTATATAG
flsynbt.fin   .CAGC..C..G..C..CC.G..C..C..C..G.....CAGC.....C..G..G..C..C.
bssyn         .CAGC..C..G..C..CC.G..C..C..C..G.....CAGC.....C..G..G..C..C.
```

Fig. 1D

```
                  1800      1810      1820      1830      1840      1850
                     *         *         *         *         *         *
BTHKURHD      ATCGAATTGAATTTGTTCCGGCAGAAGTAACCTTTGAGGCAGAATATGATTTAGAAAGAG
flsynbt.fin   .C..C..C..G..C..G..C..C..G..G.....C.....C..G..C..CC.G..G..G.
bssyn         .C..C..C..G..C..G..C..C..G..G.....C.....C..G..C..CC.G..G..G.

                  1860      1870      1880      1890      1900      1910
                     *         *         *         *         *         *
BTHKURHD      CACAAAAGGCGGTGAATGAGCTGTTTACTTCTTCCAATCAAATCGGGTTAAAAACAGATG
flsynbt.fin   .T..G.....C.....C........C..CAGCAG...C..G.....CC.G..G..C..C.
bssyn         .T..G.....C.....C........C..CAGCAG...C..G.....CC.G..G..C..C.

                  1920      1930      1940      1950      1960      1970
                     *         *         *         *         *         *
BTHKURHD      TGACGGATTATCATATTGATCAAGTATCCAATTTAGTTGAGTGTTTATCTGATGAATTTT
flsynbt.fin   ....C..C..C..C..C.....G..GAG...CC.G..G.....CC.GAGC..C..G..C.
bssyn         ....C..C..C..C..C....---------------------------------------

                  1980      1990      2000      2010      2020      2030
                     *         *         *         *         *         *
BTHKURHD      GTCTGGATGAAAAAAAAGAATTGTCCGAGAAAGTCAAACATGCGAAGCGACTTAGTGATG
flsynbt.fin   .C.....C..G..G..G..GC..AG......G..G..G..C..C.....C..G..C..C.
bssyn         ------------------------------------------------------------

                  2040      2050      2060      2070      2080      2090
                     *         *         *         *         *         *
BTHKURHD      AGCGGAATTTACTTCAAGATCCAAACTTTAGAGGGATCAATAGACAACTAGACCGTGGCT
flsynbt.fin   ....C..CC.G..G..G..C..C.....CC.C..C.....CC.C..G..G.....C....
bssyn         ------------------------------------------------------------

                  2100      2110      2120      2130      2140      2150
                     *         *         *         *         *         *
BTHKURHD      GGAGAGGAAGTACGGATATTACCATCCAAGGAGGCGATGACGTATTCAAAGAGAATTACG
flsynbt.fin   ..C.C..C..C..C..C..C........G..C.....C.....G.....G.....C....
bssyn         ------------------------------------------------------------

                  2160      2170      2180      2190      2200      2210
                     *         *         *         *         *         *
BTHKURHD      TTACGCTATTGGGTACCTTTGATGAGTGCTATCCAACGTATTTATATCAAAAAATAGATG
flsynbt.fin   .G..C..GC....C.....C..C........C..C..C..CC.G..C..G..G..C..C.
bssyn         ------------------------------------------------------------

                  2220      2230      2240      2250      2260      2270
                     *         *         *         *         *         *
BTHKURHD      AGTCGAAATTAAAAGCCTATACCCGTTACCAATTAAGAGGGTATATCGAAGATAGTCAAG
flsynbt.fin   ..AGC..GC.G..G.....C.....C.....GC.GC.C..C..C.....G..C..C..G.
bssyn         ------------------------------------------------------------

                  2280      2290      2300      2310      2320      2330
                     *         *         *         *         *         *
BTHKURHD      ACTTAGAAATCTATTTAATTCGCTACAATGCCAAACACGAAACAGTAAATGTGCCAGGTA
flsynbt.fin   ..C.G..G.....CC.G..C........C.....G.....G..C..G..C.....C..C.
bssyn         ------------------------------------------------------------

                  2340      2350      2360      2370      2380      2390
                     *         *         *         *         *         *
BTHKURHD      CGGGTTCCTTATGGCCGCTTTCAGCCCCAAGTCCAATCGGAAAATGTGCCCATCATTCCC
flsynbt.fin   .C..CAG.C.G.....C..GAGC.....C..C..C.....C..G..C.....C..CAG..
bssyn         ------------------------------------------------------------
```

Fig. 1E

```
                 2400      2410      2420      2430      2440      2450
                    *         *         *         *         *         *
BTHKURHD      ATCATTTCTCCTTGGACATTGATGTTGGATGTACAGACTTAAATGAGGACTTAGGTGTAT
flsynbt.fin   .C..C...AG.C.......C..C..G..C..C..C...C.G..C......C.G..C..G.
bssyn         ------------------------------------------------------------

                 2460      2470      2480      2490      2500      2510
                    *         *         *         *         *         *
BTHKURHD      GGGTGATATTCAAGATTAAGACGCAAGATGGCCATGCAAGACTAGGAAATCTAGAATTTC
flsynbt.fin   .......C........C.....C..G..C.....C..CC.C..G..C..C..G..G..C.
bssyn         ------------------------------------------------------------

                 2520      2530      2540      2550      2560      2570
                    *         *         *         *         *         *
BTHKURHD      TCGAAGAGAAACCATTAGTAGGAGAAGCACTAGCTCGTGTGAAAAGAGCGGAGAAAAAAT
flsynbt.fin   .G..G.....G..CC.G..G..C..G..C..G..C..C.....GC.C..C.....G..G.
bssyn         ------------------------------------------------------------

                 2580      2590      2600      2610      2620      2630
                    *         *         *         *         *         *
BTHKURHD      GGAGAGACAAACGTGAAAAATTGGAATGGGAAACAAATATTGTTTATAAAGAGGCAAAAG
flsynbt.fin   ..C.C.....G..C..G..GC....G.....G..C..C..C..G..C..G.....C..G.
bssyn         ------------------------------------------------------------

                 2640      2650      2660      2670      2680      2690
                    *         *         *         *         *         *
BTHKURHD      AATCTGTAGATGCTTTATTTGTAAACTCTCAATATGATAGATTACAAGCGGATACCAACA
flsynbt.fin   .GAGC..G..C..CC.G..C..G...AGC..G..C..CC.CC.G..G..C..C.......
bssyn         ------------------------------------------------------------

                 2700      2710      2720      2730      2740      2750
                    *         *         *         *         *         *
BTHKURHD      TCGCGATGATTCATGCGGCAGATAAACGCGTTCATAGCATTCGAGAAGCTTATCTGCCTG
flsynbt.fin   ....C.....C..C..C..C..C..G.....G..C........C..G..C..C.....C.
bssyn         ------------------------------------------------------------

                 2760      2770      2780      2790      2800      2810
                    *         *         *         *         *         *
BTHKURHD      AGCTGTCTGTGATTCCGGGTGTCAATGCGGCTATTTTTGAAGAATTAGAAGGGCGTATTT
flsynbt.fin   .....AGC.....C..C..C..G..C..C..C..C..C..G..GC.G..G..C..C..C.
bssyn         ------------------------------------------------------------

                 2820      2830      2840      2850      2860      2870
                    *         *         *         *         *         *
BTHKURHD      TCACTGCATTCTCCCTATATGATGCGAGAAATGTCATTAAAAATGGTGATTTTAATAATG
flsynbt.fin   ....C..C...AG...G..C..C..CC.C..C..G..C..G..C..C..C..C..C..C.
bssyn         ------------------------------------------------------------

                 2880      2890      2900      2910      2920      2930
                    *         *         *         *         *         *
BTHKURHD      GCTTATCCTGCTGGAACGTGAAAGGGCATGTAGATGTAGAAGAACAAAACAACCACCGTT
flsynbt.fin   ..C.GAG...............G..C..C..G..C..G..G..G..G...........CA
bssyn         ------------------------------------------------------------

                 2940      2950      2960      2970      2980      2990
                    *         *         *         *         *         *
BTHKURHD      CGGTCCTTGTTGTTCCGGAATGGGAAGCAGAAGTGTCACAAGAAGTTCGTGTCTGTCCGG
flsynbt.fin   GC..G..G..G..G..C..G.....G..C..G...AGC..G..G..G..C..G..C..C.
bssyn         ------------------------------------------------------------
```

Fig. 1F

```
                3000      3010      3020      3030      3040      3050
                   *         *         *         *         *         *
BTHKURHD      GTCGTGGCTATATCCTTCGTGTCACAGCGTACAAGGAGGGATATGGAGAAGGTTGCGTAA
flsynbt.fin   .C..C.....C.....G..C..G..C..C...........C..C..C..G..C.....G.
bssyn         ------------------------------------------------------------

                3060      3070      3080      3090      3100      3110
                   *         *         *         *         *         *
BTHKURHD      CCATTCATGAGATCGAGAACAATACAGACGAACTGAAGTTTAGCAACTGTGTAGAAGAGG
flsynbt.fin   ....C..C..............C..C.....G..C.....C........C..G..G....
bssyn         ------------------------------------------------------------

                3120      3130      3140      3150      3160      3170
                   *         *         *         *         *         *
BTHKURHD      AAGTATATCCAAACAACACGGTAACGTGTAATGATTATACTGCGACTCAAGAAGAATATG
flsynbt.fin   .G..G..C..C........C..G..C..C..C..C..C..C..C..C..G..G..G..C.
bssyn         ------------------------------------------------------------

                3180      3190      3200      3210      3220      3230
                   *         *         *         *         *         *
BTHKURHD      AGGGTACGTACACTTCTCGTAATCGAGGATATGACGGAGCCTATGAAAGCAATTCTTCTG
flsynbt.fin   ....C..C.....CAGC..C..C..C..C..C.....C.....C..G.....CAGCAGC.
bssyn         ------------------------------------------------------------

                3240      3250      3260      3270      3280      3290
                   *         *         *         *         *         *
BTHKURHD      TACCAGCTGATTATGCATCAGCCTATGAAGAAAAAGCATATACAGATGGACGAAGAGACA
flsynbt.fin   .G..C..C..C..C..CAGC.....C..G..G..G..C..C..C..C..C..CC.C....
bssyn         ------------------------------------------------------------

                3300      3310      3320      3330      3340      3350
                   *         *         *         *         *         *
BTHKURHD      ATCCTTGTGAATCTAACAGAGGATATGGGGATTACACACCACTACCAGCTGGCTATGTGA
flsynbt.fin   .C..C..C..GAGC...C.C..C..C..C..C.....C..C..G..C..C.....C....
bssyn         ------------------------------------------------------------

                3360      3370      3380      3390      3400      3410
                   *         *         *         *         *         *
BTHKURHD      CAAAAGAATTAGAGTACTTCCCAGAAACCGATAAGGTATGGATTGAGATCGGAGAAACGG
flsynbt.fin   .C..G..GC.G...........C..G.....C.....G.....C.........C..G..C.
bssyn         ------------------------------------------------------------

                3420      3430      3440      3450      3460
                   *         *         *         *         *
BTHKURHD      AAGGAACATTCATCGTGGACAGCGTGGAATTACTTCTTATGGAGGAATAA
flsynbt.fin   .G..C..C.....................GC.G..G..G.......G..G
bssyn         ------------------------------------------...TG..G
```

Fig. 2A

```
                         10        20        30        40        50        60
                          *         *         *         *         *         *
BTHKURHD         ATGGATAACAATCCGAACATCAATGAATGCATTCCTTATAATTGTTTAAGTAACCCTGAA
bssyn            .....C.....C..C........C..G.....C..C..C..C..CC.G..C.....C..G

                         70        80        90       100       110       120
                          *         *         *         *         *         *
BTHKURHD         GTAGAAGTATTAGGTGGAGAAAGAATAGAAACTGGTTACACCCCAATCGATATTTCCTTG
bssyn            ..G..G..GC.G..C..C..GC.C..C..G..C..C........C.....C..CAG.C..

                        130       140       150       160       170       180
                          *         *         *         *         *         *
BTHKURHD         TCGCTAACGCAATTTCTTTTGAGTGAATTTGTTCCCGGTGCTGGATTTGTGTTAGGACTA
bssyn            AGC..G..C..G..C..GC....C..G..C..G.....C..C..C..C...C.G..C..G

                        190       200       210       220       230       240
                          *         *         *         *         *         *
BTHKURHD         GTTGATATAATATGGGGAATTTTTGGTCCCTCTCAATGGGACGCATTTCTTGTACAAATT
bssyn            ..G..C..C..C.....C..C..C..C...AGC..G........C..C..G..G..G..C

                        250       260       270       280       290       300
                          *         *         *         *         *         *
BTHKURHD         GAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAGGAACCAAGCCATTTCTAGATTA
bssyn            ..G...C.G..C.....GC.C..C..G..G.....CC.C.....G.....CAGCC.CC.G

                        310       320       330       340       350       360
                          *         *         *         *         *         *
BTHKURHD         GAAGGACTAAGCAATCTTTATCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAGAT
bssyn            ..G..C..G.....C..G..C.....C.....C..GAGC..CC.C........G..C..C

                        370       380       390       400       410       420
                          *         *         *         *         *         *
BTHKURHD         CCTACTAATCCAGCATTAAGAGAAGAGATGCGTATTCAATTCAATGACATGAACAGTGCC
bssyn            ..C..C..C..C..CC.GC.C..G........C..C..G.....C...........C...

                        430       440       450       460       470       480
                          *         *         *         *         *         *
BTHKURHD         CTTACAACCGCTATTCCTCTTTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTA
bssyn            ..G..C.....C..C..C..G..C..C..G..G..C..C..G..G..C..GC.GAGC..G

                        490       500       510       520       530       540
                          *         *         *         *         *         *
BTHKURHD         TATGTTCAAGCTGCAAATTTACATTTATCAGTTTTGAGAGATGTTTCAGTGTTTGGACAA
bssyn            ..C..G..G..C..C..CC.G..CC.GAGC..GC..C.C..C..CAGC.....C..C..G

                        550       560       570       580       590       600
                          *         *         *         *         *         *
BTHKURHD         AGGTGGGGATTTGATGCCGCGACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT
bssyn            C.C.....C..C..C.....C..C.....C..C..C..C..C..CC.G..CC.C..G..C

                        610       620       630       640       650       660
                          *         *         *         *         *         *
BTHKURHD         GGCAACTATACAGATCATGCTGTACGCTGGTACAATACGGGATTAGAGCGTGTATGGGGA
bssyn            ........C..C..C..C..C..G...........C..C..CC.G.....C..G.....T

                        670       680       690       700       710       720
                          *         *         *         *         *         *
BTHKURHD         CCGGATTCTAGAGATTGGATAAGATATAATCAATTTAGAAGAGAATTAACACTAACTGTA
bssyn            ..C..CAGCC.C..C.....C..G..C..C..G..CC.CC.C..GC.G..C..G..C..G
```

Fig. 2B

```
                    730       740       750       760       770       780
                      *         *         *         *         *         *
BTHKURHD   TTAGATATCGTTTCTCTATTTCCGAACTATGATAGTAGAACGTATCCAATTCGAACAGTT
bssyn      C.G..C.....GAGC..G..C..C.....C..C..CC.C..C..C..C..C..C..C..G

                    790       800       810       820       830       840
                      *         *         *         *         *         *
BTHKURHD   TCCCAATTAACAAGAGAAATTTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTTTT
bssyn      AG...GC.G..CC.C..G.....C..C.....C..GC.G..G..C..C..C..C..C..C

                    850       860       870       880       890       900
                      *         *         *         *         *         *
BTHKURHD   CGAGGCTCGGCTCAGGGCATAGAAGGAAGTATTAGGAGTCCACATTTGATGGATATACTT
bssyn      ..C...AGC..C........C..G..C..C..CC.C..C..C..CC.......C..C..G

                    910       920       930       940       950       960
                      *         *         *         *         *         *
BTHKURHD   AACAGTATAACCATCTATACGGATGCTCATAGAGGAGAATATTATTGGTCAGGGCATCAA
bssyn      .....C..C........C..C..C..C..CC.C..C..G..C..C...AGC..C..C..G

                    970       980       990      1000      1010      1020
                      *         *         *         *         *         *
BTHKURHD   ATAATGGCTTCTCCTGTAGGGTTTTCGGGGCCAGAATTCACTTTTCCGCTATATGGAACT
bssyn      ..C.....CAGC..C..C..C..CAGC..C..C..G.....C..C..C..G..C..C..C

                   1030      1040      1050      1060      1070      1080
                      *         *         *         *         *         *
BTHKURHD   ATGGGAAATGCAGCTCCACAACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATAGA
bssyn      .....C..C..T..A..T..G..G..C..C..G..A..G..G..C.....A.....CC.C

                   1090      1100      1110      1120      1130      1140
                      *         *         *         *         *         *
BTHKURHD   ACATTATCGTCCACTTTATATAGAAGACCTTTTAATATAGGGATAAATAATCAACAACTA
bssyn      ..CC.GAGCAG...CC.G..CC.TC.......C..C..C..C..C..C..C..G..G..G

                   1150      1160      1170      1180      1190      1200
                      *         *         *         *         *         *
BTHKURHD   TCTGTTCTTGACGGGACAGAATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTA
bssyn      AGC..G..G.....C..C..G..C..C..C..C...AG.AGC..CC....CAG...C..G

                   1210      1220      1230      1240      1250      1260
                      *         *         *         *         *         *
BTHKURHD   TACAGAAAAAGCGGAACGGTAGATTCGCTGGATGAAATACCGCCACAGAATAACAACGTG
bssyn      ...C.C..G.....C..C..G..CAGC.....C..G..C..C..T.....C.........

                   1270      1280      1290      1300      1310      1320
                      *         *         *         *         *         *
BTHKURHD   CCACCTAGGCAAGGATTTAGTCATCGATTAAGCCATGTTTCAATGTTTCGTTCAGGCTTT
bssyn      ......C.A..G..C..C..C..C..TC.G.....C..GAGC.....C..CAGT.....C

                   1330      1340      1350      1360      1370      1380
                      *         *         *         *         *         *
BTHKURHD   AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCTCTTGGATACATCGTAGTGCT
bssyn      ..C..C..C..C..G..C..C..CC.T..A.........AGC.....T..C..C.....C

                   1390      1400        1410      1420      1430
                      *         *           *         *         *
BTHKURHD   GAATTTAATAATATAATTCCTTCATCA--CAAATTACACAAATACCTTTAACAAAATCTA
bssyn      ..G..C..C..C..C..C..--..G..GC..G..C..C..G..C..CC.G..C..GAGC.
```

Fig. 2C

```
                1440      1450      1460      1470      1480      1490
                   *         *         *         *         *         *
BTHKURHD      CTAATCTTGGCTCTGGAACTTCTGTCGTTAAAGGACCAGGATTTACAGGAGGAGATATTC
bssyn         .C..C..G...AGC..C..CAGC..G..G..G..C..C..C..C..C..C..C..C..C.

                1500      1510      1520      1530      1540      1550
                   *         *         *         *         *         *
BTHKURHD      TTCGAAGAACTTCACCTGGCCAGATTTCAACCTTAAGAGTAAATATTACTGCACCATTAT
bssyn         .G..CC.C..CAGC..C........CAGC...C.GC.C..G..C..C..C..C..CC.GA

                1560      1570      1580      1590      1600      1610
                   *         *         *         *         *         *
BTHKURHD      CACAAAGATATCGGGTAAGAATTCGCTACGCTTCTACCACAAATTTACAATTCCATACAT
bssyn         GC..GC.C..C..C..CC.C..C........CAGC.....C..CC.G..G.....C..CA

                1620      1630      1640      1650      1660      1670
                   *         *         *         *         *         *
BTHKURHD      CAATTGACGGAAGACCTATTAATCAGGGGAATTTTTCAGCAACTATGAGTAGTGGGAGTA
bssyn         GC..C.....CC.C..C..C..C.....C..C..CAGC..C..C.....C..C..C..C.

                1680      1690      1700      1710      1720      1730
                   *         *         *         *         *         *
BTHKURHD      ATTTACAGTCCGGAAGCTTTAGGACTGTAGGTTTTACTACTCCGTTTAACTTTTCAAATG
bssyn         .CC.G...AG...C.....CC.C..C..G..C..C..C..C..C..C.....CAGC..C.

                1740      1750      1760      1770      1780      1790
                   *         *         *         *         *         *
BTHKURHD      GATCAAGTGTATTTACGTTAAGTGCTCATGTCTTCAATTCAGGCAATGAAGTTTATATAG
bssyn         .CAGC..C..G..C..CC.G..C..C..C..G.....CAGC.....C..G..G..C..C.

                1800      1810      1820      1830      1840      1850
                   *         *         *         *         *         *
BTHKURHD      ATCGAATTGAATTTGTTCCGGCAGAAGTAACCTTTGAGGCAGAATATGATTTAGAAAGAG
bssyn         .C..C..C..G..C..G..C..C..G..G.....C.....C..G..C..CC.G..G..G.

                1860      1870      1880      1890      1900      1910
                   *         *         *         *         *         *
BTHKURHD      CACAAAAGGCGGTGAATGAGCTGTTTACTTCTTCCAATCAAATCGGGTTAAAAACAGATG
bssyn         .T..G.....C.....C........C..CAGCAG...C..G.....CC.G..G..C..C.

                1920      1930      1940
                   *         *         *
BTHKURHD      TGACGGATTATCATATTGATCAAGTATCC
bssyn         ....C..C..C..C..C.....G..G.AG
```

Fig. 3A

```
                    10         20         30         40         50         60
                     *          *          *          *          *          *
syn1T.mze    ATGGACAACAACCCCAACATCAACGAGTGCATCCCCTACAACTGCCTGAGCAACCCCGAG
bssyn        ............................................................
synful.mod   ............................................................

                    70         80         90        100        110        120
                     *          *          *          *          *          *
syn1T.mze    GTGGAGGTGCTGGGCGGCGAGCGCATCGAGACCGGCTACACCCCCATCGACATCAGCCTG
bssyn        ............................................................
synful.mod   ............................................................

                   130        140        150        160        170        180
                     *          *          *          *          *          *
syn1T.mze    AGCCTGACCCAGTTCCTGCTGAGCGAGTTCGTGCCCGGCGCCGGCTTCGTGCTGGGCCTG
bssyn        ............................................................
synful.mod   ............................................................

                   190        200        210        220        230        240
                     *          *          *          *          *          *
syn1T.mze    GTGGACATCATCTGGGGCATCTTCGGCCCCAGCCAGTGGGACGCCTTCCTGGTGCAGATC
bssyn        ............................................................
synful.mod   ............................................................

                   250        260        270        280        290        300
                     *          *          *          *          *          *
syn1T.mze    GAGCAGCTGATCAACCAGCGCATCGAGGAGTTCGCCCGCAACCAGGCCATCAGCCGCCTG
bssyn        ............................................................
synful.mod   ............................................................

                   310        320        330        340        350        360
                     *          *          *          *          *          *
syn1T.mze    GAGGGCCTGAGCAACCTGTACCAGATCTACGCCGAGAGCTTCCGCGAGTGGGAGGCCGAC
bssyn        .......................A....................................
synful.mod   .......................A....................................

                   370        380        390        400        410        420
                     *          *          *          *          *          *
syn1T.mze    CCCACCAACCCCGCCCTGCGCGAGGAGATGCGCATCCAGTTCAACGACATGAACAGCGCC
bssyn        ............................................................
synful.mod   ............................................................

                   430        440        450        460        470        480
                     *          *          *          *          *          *
syn1T.mze    CTGACCACCGCCATCCCCCTGTTCGCCGTGCAGAACTACCAGGTGCCCCTGCTGAGCGTG
bssyn        ............................................................
synful.mod   ............................................................

                   490        500        510        520        530        540
                     *          *          *          *          *          *
syn1T.mze    TACGTGCAGGCCGCCAACCTGCACCTGAGCGTGCTGCGCGACGTGAGCGTGTTCGGCCAG
bssyn        ............................................C...............
synful.mod   ............................................C...............

                   550        560        570        580        590        600
                     *          *          *          *          *          *
syn1T.mze    CGCTGGGGCTTCGACGCCGCCACCATCAACAGCCGCTACAACGACCTGACCCGCCTGATC
bssyn        ............................................................
synful.mod   ............................................................
```

Fig. 3B

```
                       610       620       630       640       650       660
                         *         *         *         *         *         *
syn1T.mze       GGCAACTACACCGACCACGCCGTGCGCTGGTACAACACCGGCCTGGAGCGCGTGTGGGGC
bssyn           ...........................................................T
synful.mod      ...........................................................T

                       670       680       690       700       710       720
                         *         *         *         *         *         *
syn1T.mze       CCCGACAGCCGCGACTGGATCCGCTACAACCAGTTCCGCCGCGAGCTGACCCTGACCGTG
bssyn           .....................A.G....................................
synful.mod      .....................A.G....................................

                       730       740       750       760       770       780
                         *         *         *         *         *         *
syn1T.mze       CTGGACATCGTGAGCCTGTTCCCCAACTACGACAGCCGCACCTACCCCATCCGCACCGTG
bssyn           ............................................................
synful.mod      ............................................................

                       790       800       810       820       830       840
                         *         *         *         *         *         *
syn1T.mze       AGCCAGCTGACCCGCGAGATCTACACCAACCCCGTGCTGGAGAACTTCGACGGCAGCTTC
bssyn           ....................T.......................................
synful.mod      ....................T.......................................

                       850       860       870       880       890       900
                         *         *         *         *         *         *
syn1T.mze       CGCGGCAGCGCCCAGGGCATCGAGGGCAGCATCCGCAGCCCCCACCTGATGGACATCCTG
bssyn           ............................................................
synful.mod      ............................................................

                       910       920       930       940       950       960
                         *         *         *         *         *         *
syn1T.mze       AACAGCATCACCATCTACACCGACGCCCACCGCGGCGAGTACTACTGGAGCGGCCACCAG
bssyn           ............................................................
synful.mod      ............................................................

                       970       980       990      1000      1010      1020
                         *         *         *         *         *         *
syn1T.mze       ATCATGGCCAGCCCCGTGGGCTTCAGCGGCCCCGAGTTCACCTTCCCCCTGTACGGCACC
bssyn           .................C..........................................
synful.mod      .................C..........................................

                      1030      1040      1050      1060      1070      1080
                         *         *         *         *         *         *
syn1T.mze       ATGGGCAACGCCGCCCCCCAGCAGCGCATCGTGGCCCAGCTGGGCCAGGGCGTGTACCGC
bssyn           ...........T..A..T.................A..............A.........
synful.mod      ...........T..A..T.................A..............A.........

                      1090      1100      1110      1120      1130      1140
                         *         *         *         *         *         *
syn1T.mze       ACCCTGAGCAGCACCCTGTACCGCCGCCCCTTCAACATCGGCATCAACAACCAGCAGCTG
bssyn           .......................T..A..T..............................
synful.mod      .......................T..A..T..............................

                      1150      1160      1170      1180      1190      1200
                         *         *         *         *         *         *
syn1T.mze       AGCGTGCTGGACGGCACCGAGTTCGCCTACGGCACCAGCAGCAACCTGCCCAGCGCCGTG
bssyn           ............................................................
synful.mod      ............................................................
```

Fig. 3C

```
                      1210      1220      1230      1240      1250      1260
                         *         *         *         *         *         *
syn1T.mze   TACCGCAAGAGCGGCACCGTGGACAGCCTGGACGAGATCCCCCCCCAGAACAACAACGTG
bssyn       ............................................T...............
synful.mod  ............................................T...............

                      1270      1280      1290      1300      1310      1320
                         *         *         *         *         *         *
syn1T.mze   CCCCCCCGCCAGGGCTTCAGCCACCGCCTGAGCCACGTGAGCATGTTCCGCAGCGGCTTC
bssyn       ..A..T..A.................T..........................T......
synful.mod  ..A..T..A.................T..........................T......

                      1330      1340      1350      1360      1370      1380
                         *         *         *         *         *         *
syn1T.mze   AGCAACAGCAGCGTGAGCATCATCCGCGCCCCCATGTTCAGCTGGATCCACCGCAGCGCC
bssyn       ..........................T..A..T..............T........T...
synful.mod  ..........................T..A..T..............T........T...

                      1390      1400      1410      1420      1430      1440
                         *         *         *         *         *         *
syn1T.mze   GAGTTCAACAACATCATCCCCAGCAGCCAGATCACCCAGATCCCCCTGACCAAGAGCACC
bssyn       ............................................................
synful.mod  ............................................................

                      1450      1460      1470      1480      1490      1500
                         *         *         *         *         *         *
syn1T.mze   AACCTGGGCAGCGGCACCAGCGTGGTGAAGGGCCCCGGCTTCACCGGCGGCGACATCCTG
bssyn       ............................................................
synful.mod  ............................................................

                      1510      1520      1530      1540      1550      1560
                         *         *         *         *         *         *
syn1T.mze   CGCCGCACCAGCCCCGGCCAGATCAGCACCCTGCGCGTGAACATCACCGCCCCCCTGAGC
bssyn       ............................................................
synful.mod  ............................................................

                      1570      1580      1590      1600      1610      1620
                         *         *         *         *         *         *
syn1T.mze   CAGCGCTACCGCGTGCGCATCCGCTACGCCAGCACCACCAACCTGCAGTTCCACACCAGC
bssyn       ..............C.............................................
synful.mod  ..............C.............................................

                      1630      1640      1650      1660      1670      1680
                         *         *         *         *         *         *
syn1T.mze   ATCGACGGCCGCCCCATCAACCAGGGCAACTTCAGCGCCACCATGAGCAGCGGCAGCAAC
bssyn       ............................................................
synful.mod  ............................................................

                      1690      1700      1710      1720      1730      1740
                         *         *         *         *         *         *
syn1T.mze   CTGCAGAGCGGCAGCTTCCGCACCGTGGGCTTCACCACCCCCTTCAACTTCAGCAACGGC
bssyn       ............................................................
synful.mod  ............................................................

                      1750      1760      1770      1780      1790      1800
                         *         *         *         *         *         *
syn1T.mze   AGCAGCGTGTTCACCCTGAGCGCCCACGTGTTCAACAGCGGCAACGAGGTGTACATCGAC
bssyn       ............................................................
synful.mod  ............................................................
```

Fig. 3D

```
                   1810      1820      1830      1840      1850      1860
                      *         *         *         *         *         *
syn1T.mze    CGCATCGAGTTCGTGCCCGCCGAGGTGACCTTCGAGGCCGAGTACGACCTGGAGCGCGCC
bssyn        ......................................................A.G..T
synful.mod   ......................................................A.G..T

                   1870      1880      1890      1900      1910      1920
                      *         *         *         *         *         *
syn1T.mze    CAGAAGGCCGTGAACGAGCTGTTCACCAGCAGCAACCAGATCGGCCTGAAGACCGACGTG
bssyn        ............................................................
synful.mod   ............................................................

                   1930      1940      1950      1960      1970      1980
                      *         *         *         *         *         *
syn1T.mze    ACCGACTACCACATCGACCAGGTGAGCAACCTGGTGGAGTGCCTGAGCGACGAGTTCTGC
bssyn        .................T.....-------------------------------------
synful.mod   .................T..........................................

                   1990      2000      2010      2020      2030      2040
                      *         *         *         *         *         *
syn1T.mze    CTGGACGAGAAGAAGGAGCTGAGCGAGAAGGTGAAGCACGCCAAGCGCCTGAGCGACGAG
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2050      2060      2070      2080      2090      2100
                      *         *         *         *         *         *
syn1T.mze    CGCAACCTGCTGCAGGACCCCAACTTCCGCGGCATCAACCGCCAGCTGGACCGCGGCTGG
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2110      2120      2130      2140      2150      2160
                      *         *         *         *         *         *
syn1T.mze    CGCGGCAGCACCGACATCACCATCCAGGGCGGCGACGACGTGTTCAAGGAGAACTACGTG
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2170      2180      2190      2200      2210      2220
                      *         *         *         *         *         *
syn1T.mze    ACCCTGCTGGGCACCTTCGACGAGTGCTACCCCACCTACCTGTACCAGAAGATCGACGAG
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2230      2240      2250      2260      2270      2280
                      *         *         *         *         *         *
syn1T.mze    AGCAAGCTGAAGGCCTACACCCGCTACCAGCTGCGCGGCTACATCGAGGACAGCCAGGAC
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2290      2300      2310      2320      2330      2340
                      *         *         *         *         *         *
syn1T.mze    CTGGAGATCTACCTGATCCGCTACAACGCCAAGCACGAGACCGTGAACGTGCCCGGCACC
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2350      2360      2370      2380      2390      2400
                      *         *         *         *         *         *
syn1T.mze    GGCAGCCTGTGGCCCCTGAGCGCCCCCAGCCCCATCGGCAAGTGCGCCCACCACAGCCAC
bssyn        ------------------------------------------------------------
synful.mod   ............................................................
```

Fig. 3E

```
                   2410      2420      2430      2440      2450      2460
                      *         *         *         *         *         *
syn1T.mze    CACTTCAGCCTGGACATCGACGTGGGCTGCACCGACCTGAACGAGGACCTGGGCGTGTGG
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2470      2480      2490      2500      2510      2520
                      *         *         *         *         *         *
syn1T.mze    GTGATCTTCAAGATCAAGACCCAGGACGGCCACGCCCGCCTGGGCAACCTGGAGTTCCTG
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2530      2540      2550      2560      2570      2580
                      *         *         *         *         *         *
syn1T.mze    GAGGAGAAGCCCCTGGTGGGCGAGGCCCTGGCCCGCGTGAAGCGCGCCGAGAAGAAGTGG
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2590      2600      2610      2620      2630      2640
                      *         *         *         *         *         *
syn1T.mze    CGCGACAAGCGCGAGAAGCTGGAGTGGGAGACCAACATCGTGTACAAGGAGGCCAAGGAG
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2650      2660      2670      2680      2690      2700
                      *         *         *         *         *         *
syn1T.mze    AGCGTGGACGCCCTGTTCGTGAACAGCCAGTACGACCGCCTGCAGGCCGACACCAACATC
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2710      2720      2730      2740      2750      2760
                      *         *         *         *         *         *
syn1T.mze    GCCATGATCCACGCCGCCGACAAGCGCGTGCACAGCATCCGCGAGGCCTACCTGCCCGAG
bssyn        ------------------------------------------------------------
synful.mod   ......................;...............T.....................

                   2770      2780      2790      2800      2810      2820
                      *         *         *         *         *         *
syn1T.mze    CTGAGCGTGATCCCCGGCGTGAACGCCGCCATCTTCGAGGAGCTGGAGGGCCGCATCTTC
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2830      2840      2850      2860      2870      2880
                      *         *         *         *         *         *
syn1T.mze    ACCGCCTTCAGCCTGTACGACGCCCGCAACGTGATCAAGAACGGCGACTTCAACAACGGC
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2890      2900      2910      2920      2930      2940
                      *         *         *         *         *         *
syn1T.mze    CTGAGCTGCTGGAACGTGAAGGGCCACGTGGACGTGGAGGAGCAGAACAACCACCGCAGC
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                   2950      2960      2970      2980      2990      3000
                      *         *         *         *         *         *
syn1T.mze    GTGCTGGTGGTGCCCGAGTGGGAGGCCGAGGTGAGCCAGGAGGTGCGCGTGTGCCCCGGC
bssyn        ------------------------------------------------------------
synful.mod   ............................................................
```

Fig. 3F

```
                    3010      3020      3030      3040      3050      3060
                       *         *         *         *         *         *
syn1T.mze    CGCGGCTACATCCTGCGCGTGACCGCCTACAAGGAGGGCTACGGCGAGGGCTGCGTGACC
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                    3070      3080      3090      3100      3110      3120
                       *         *         *         *         *         *
syn1T.mze    ATCCACGAGATCGAGAACAACACCGACGAGCTGAAGTTCAGCAACTGCGTGGAGGAGGAG
bssyn        ------------------------------------------------------------
synful.mod   ................................C...........................

                    3130      3140      3150      3160      3170      3180
                       *         *         *         *         *         *
syn1T.mze    GTGTACCCCAACAACACCGTGACCTGCAACGACTACACCGCCACCCAGGAGGAGTACGAG
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                    3190      3200      3210      3220      3230      3240
                       *         *         *         *         *         *
syn1T.mze    GGCACCTACACCAGCCGCAACCGCGGCTACGACGGCGCCTACGAGAGCAACAGCAGCGTG
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                    3250      3260      3270      3280      3290      3300
                       *         *         *         *         *         *
syn1T.mze    CCCGCCGACTACGCCAGCGCCTACGAGGAGAAGGCCTACACCGACGGCCGCCGCGACAAC
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                    3310      3320      3330      3340      3350      3360
                       *         *         *         *         *         *
syn1T.mze    CCCTGCGAGAGCAACCGCGGCTACGGCGACTACACCCCCCTGCCCGCCGGCTACGTGACC
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                    3370      3380      3390      3400      3410      3420
                       *         *         *         *         *         *
syn1T.mze    AAGGAGCTGGAGTACTTCCCCGAGACCGACAAGGTGTGGATCGAGATCGGCGAGACCGAG
bssyn        ------------------------------------------------------------
synful.mod   ............................................................

                    3430      3440      3450      3460
                       *         *         *         *
syn1T.mze    GGCACCTTCATCGTGGACAGCGTGGAGCTGCTGCTGATGGAGGAGTAG
bssyn        --------------------------------------------....
synful.mod   ................................................
```

Fig. 4A

```
                  10        20        30        40        50        60
                   *         *         *         *         *         *
BTHKURHD  ATGGATAACAATCCGAACATCAATGAATGCATTCCTTATAATTGTTTAAGTAACCCTGAA
PMONBT    .....C.....C..A........C...........A..C..C..C..G........A...
bssyn     .....C.....C..C........C..G.....C..C..C..C..CC.G..C.....C..G

                  70        80        90       100       110       120
                   *         *         *         *         *         *
BTHKURHD  GTAGAAGTATTAGGTGGAGAAAGAATAGAAACTGGTTACACCCCAATCGATATTTCCTTG
PMONBT    ..T......C.T.........C.C..T.....C.........T..C.....C..C......
bssyn     ..G..G..GC.G..C..C..GC.C..C..G..C..C.........C.....C..CAG.C..

                 130       140       150       160       170       180
                   *         *         *         *         *         *
BTHKURHD  TCGCTAACGCAATTTCTTTTGAGTGAATTTGTTCCCGGTGCTGGATTTGTGTTAGGACTA
PMONBT    ..CT.G..A..G.....GC.C..C..G..C..G..A.........G..C..TC.C......
bssyn     AGC..G..C..G..C..GC....C..G..C..G.....C..C..C..C...C.G..C..G

                 190       200       210       220       230       240
                   *         *         *         *         *         *
BTHKURHD  GTTGATATAATATGGGGAATTTTTGGTCCCTCTCAATGGACGCATTTCTTGTACAAATT
PMONBT    .....C..C..C.....T..C.........A.............T.....C..G..G......
bssyn     ..G..C..C..C.....C..C..C..C...AGC..G.........C..C..G..G..G..C

                 250       260       270       280       290       300
                   *         *         *         *         *         *
BTHKURHD  GAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAGGAACCAAGCCATTTCTAGATTA
PMONBT    ..G.....G..C.....G..G..C.....G.....C.........G.....C.....G..G
bssyn     ..G...C.G..C.....GC.C..C..G..G.....CC.C.....G.....CAGCC.CC.G

                 310       320       330       340       350       360
                   *         *         *         *         *         *
BTHKURHD  GAAGGACTAAGCAATCTTTATCAAATTTACGCAGAATCTTTTAGAGAGTGGGAAGCAGAT
PMONBT    ......T.G.........C..C.....C..T.....GAGC..C...............C...
bssyn     ..G..C..G.....C..G..C......C.....C..GAGC..CC.C.........G..C..C

                 370       380       390       400       410       420
                   *         *         *         *         *         *
BTHKURHD  CCTACTAATCCAGCATTAAGAGAAGAGATGCGTATTCAATTCAATGACATGAACAGTGCC
PMONBT    .........C.....TC.CC.C..G..A..................C............C...
bssyn     ..C..C..C..C..CC.GC.C..G.........C..C..G.....C............C...

                 430       440       450       460       470       480
                   *         *         *         *         *         *
BTHKURHD  CTTACAACCGCTATTCCTCTTTTTGCAGTTCAAAATTATCAAGTTCCTCTTTTATCAGTA
PMONBT    T.G..C..A.....C..AT.G..C......C..G..C..C...........C..G..C..G
bssyn     ..G..C.....C..C..C..G..C..C..G..G..C..C..G..G..C..GC.GAGC..G

                 490       500       510       520       530       540
                   *         *         *         *         *         *
BTHKURHD  TATGTTCAAGCTGCAAATTTACATTTATCAGTTTTGAGAGATGTTTCAGTGTTTGGACAA
PMONBT    ..C.........A..T...C.T..CC.CAGC..GC.TC.....C...AGC.........G...
bssyn     ..C..G..G..C..C..CC.G..CC.GAGC..GC..C.C..C..CAGC.....C..C..G

                 550       560       570       580       590       600
                   *         *         *         *         *         *
BTHKURHD  AGGTGGGGATTTGATGCCGCGACTATCAATAGTCGTTATAATGATTTAACTAGGCTTATT
PMONBT    ............C.....T..A..C.........C.....C..C..CC.T.........G...
bssyn     C.C.....C..C..C.....C..C.....C..C..C..C..C..CC.G..CC.C..G..C
```

Fig. 4B

```
                     610       620       630       640       650       660
                       *         *         *         *         *         *
BTHKURHD   GGCAACTATACAGATCATGCTGTACGCTGGTACAATACGGATTAGAGCGTGTATGGGGA
PMONBT     ..A.....C..C..C..C.....T..T.........C..T..C..G........C.....T
bssyn      ........C..C..C..C..C..G...........C..C..CC.G.....C..G.....T

                     670       680       690       700       710       720
                       *         *         *         *         *         *
BTHKURHD   CCGGATTCTAGAGATTGGATAAGATATAATCAATTTAGAAGAGAATTAACACTAACTGTA
PMONBT     ..T.................T.....C..C..G..C..G........G..C..C..A..T
bssyn      ..C..CAGCC.C..C.....C..G..C..C..G..CC.CC.C..GC.G..C..G..C..G

                     730       740       750       760       770       780
                       *         *         *         *         *         *
BTHKURHD   TTAGATATCGTTTCTCTATTTCCGAACTATGATAGTAGAACGTATCCAATTCGAACAGTT
PMONBT     ..G..C..T..G.....C..C...........CTCC.....C..C..T..C..T.....G
bssyn      C.G..C......GAGC..G..C..C.....C..C..CC.C..C..C..C..C..C..C..G

                     790       800       810       820       830       840
                       *         *         *         *         *         *
BTHKURHD   TCCCAATTAACAAGAGAAATTTATACAAACCCAGTATTAGAAAATTTTGATGGTAGTTTT
PMONBT     ......C.T..C.........C.....T.........TC.T..G..C..C..C.....C..C
bssyn      AG...GC.G..CC.C..G.....C..C.....C..GC.G..G..C..C..C..C..C..C

                     850       860       870       880       890       900
                       *         *         *         *         *         *
BTHKURHD   CGAGGCTCGGCTCAGGGCATAGAAGGAAGTATTAGGAGTCCACATTTGATGGATATACTT
PMONBT     ..T..T..T..C..A..T..C.....CTCC..C.....C.....C........C..CT.G
bssyn      ..C...AGC..C.........C..G..C..C..CC.C..C..C..CC.......C..C..G

                     910       920       930       940       950       960
                       *         *         *         *         *         *
BTHKURHD   AACAGTATAACCATCTATACGGATGCTCATAGAGGAGAATATTATTGGTCAGGGCATCAA
PMONBT     .....C.....T.....C.GC.........C.........G.....C.....T..A..C..G
bssyn      .....C..C.........C..C..C..C..CC.C..C..G..C..C...AGC..C..C..G

                     970       980       990      1000      1010      1020
                       *         *         *         *         *         *
BTHKURHD   ATAATGGCTTCTCCTGTAGGGTTTTCGGGGCCAGAATTCACTTTTCCGCTATATGGAACT
PMONBT     ..C.....C.....A..T..A..CAGC.....C..G..T..C.....T..C..........
bssyn      ..C.....CAGC..C..C..C..CAGC..C..C..G.....C..C..C..G..C..C..C

                    1030      1040      1050      1060      1070      1080
                       *         *         *         *         *         *
BTHKURHD   ATGGGAAATGCAGCTCCACAACAACGTATTGTTGCTCAACTAGGTCAGGGCGTGTATAGA
PMONBT     ........C..C.....................C.....................T..C..C...
bssyn      .....C..C..T..A..T..G..G..C..C..G..A..G..G..C.....A.....CC.C

                    1090      1100      1110      1120      1130      1140
                       *         *         *         *         *         *
BTHKURHD   ACATTATCGTCCACTTTATATAGAAGACCTTTTAATATAGGGATAAATAATCAACAACTA
PMONBT     ..C..G..T.....C..G..C.........C..C.....C..T..C..C..C..G.....T
bssyn      ..CC.GAGCAG...CC.G..CC.TC.......C..C..C..C..C..C..C..G..G..G

                    1150      1160      1170      1180      1190      1200
                       *         *         *         *         *         *
BTHKURHD   TCTGTTCTTGACGGGACAGAATTTGCTTATGGAACCTCCTCAAATTTGCCATCCGCTGTA
PMONBT     ..C............A.....G..C..C............T..T..C...............T
bssyn      AGC..G..G.....C..C..G..C..C..C..C...AG.AGC..CC....CAG...C..G
```

Fig. 4C

```
                1210      1220      1230      1240      1250      1260
                   *         *         *         *         *         *
BTHKURHD  TACAGAAAAAGCGGAACGGTAGATTCGCTGGATGAAATACCGCCACAGAATAACAACGTG
PMONBT    ........G........C..T....CT.....C.....C..A........C.....T...
bssyn     ...C.C..G.....C..C..G..CAGC.....C..G..C..C..T.....C.........

                1270      1280      1290      1300      1310      1320
                   *         *         *         *         *         *
BTHKURHD  CCACCTAGGCAAGGATTTAGTCATCGATTAAGCCATGTTTCAATGTTTCGTTCAGGCTTT
PMONBT    .....C...........CTCC..CA.G..G.....C..G..C.....C.....C..A..C
bssyn     ......C.A..G..C..C..C..C..TC.G.....C..GAGC.....C..CAGT.....C

                1330      1340      1350      1360      1370      1380
                   *         *         *         *         *         *
BTHKURHD  AGTAATAGTAGTGTAAGTATAATAAGAGCTCCTATGTTCTCTTGGATACATCGTAGTGCT
PMONBT    ..C..C...TCC..G..C..C..C.................A.....T............
bssyn     ..C..C..C..C..G..C..C..CC.T..A.........AGC.....T..C..C.....C

                1390      1400        1410      1420      1430
                   *         *           *         *         *
BTHKURHD  GAATTTAATAATATAATTCCTTCATCA  CAAATTACACAAATACCTTTAACAAAATCTA
PMONBT    ..G..C..C.....C........C..T  .....C..C.....C..A..G..C..G....
bssyn     ..G..C..C..C..C..C..  ..G..GC..G..C..C..G..C..CC.G..C..GAGC.

       1440      1450      1460      1470      1480      1490
           *         *         *         *         *         *
BTHKURHD  CTAATCTTGGCTCTGGAACTTCTGTCGTTAAAGGACCAGGATTTACAGGAGGAGATATTC
PMONBT    ....C.....A.................G...........C..C........T.......
bssyn     .C..C..G...AGC..C..CAGC..G..G..G..C..C..C..C..C..C..C..C..C.

       1500      1510      1520      1530      1540      1550
           *         *         *         *         *         *
BTHKURHD  TTCGAAGAACTTCACCTGGCCAGATTTCAACCTTAAGAGTAAATATTACTGCACCATTAT
PMONBT    ..A..........T............AGC...C.C.....T..C..C.........C.T.
bssyn     .G..CC.C..CAGC..C........CAGC...C.GC.C..G..C..C..C..C..CC.GA

       1560      1570      1580      1590      1600      1610
           *         *         *         *         *         *
BTHKURHD  CACAAAGATATCGGGTAAGAATTCGCTACGCTTCTACCACAAATTTACAATTCCATACAT
PMONBT    .T...........T..C..G.....T.....A........T..C..G........C..C.
bssyn     GC..GC.C..C..C..CC.C..C........CAGC.....C..CC.G..G.....C..CA

       1620      1630      1640      1650      1660      1670
           *         *         *         *         *         *
BTHKURHD  CAATTGACGGAAGACCTATTAATCAGGGGAATTTTTCAGCAACTATGAGTAGTGGGAGTA
PMONBT    .C..C........G.....C........T..C..C..C.....C...TCA..C..C..C.
bssyn     GC..C.....CC.C..C..C..C.....C..C..CAGC..C..C.....C..C..C..C.

       1680      1690      1700      1710      1720      1730
           *         *         *         *         *         *
BTHKURHD  ATTTACAGTCCGGAAGCTTTAGGACTGTAGGTTTTACTACTCCGTTTAACTTTTCAAATG
PMONBT    .C..G..A.....C.....C..A..C..C.....C........T..C.....C..T..C.
bssyn     .CC.G...AG...C.....CC.C..C..G..C..C..C..C..C..C.....CAGC..C.

       1740      1750      1760      1770      1780      1790
           *         *         *         *         *         *
BTHKURHD  GATCAAGTGTATTTACGTTAAGTGCTCATGTCTTCAATTCAGGCAATGAAGTTTATATAG
PMONBT    .......C..T..C..CC.T..C........G........T...........G..C..T.
bssyn     .CAGC..C..G..C..CC.G..C..C..C..G.....CAGC.....C..G..G..C..C.
```

Fig. 4D

```
              1800      1810      1820      1830      1840
                 *         *         *         *         *
BTHKURHD      ATCGAATTGAATTTGTTCCGGCAGAAGTAACCTTTGAGGCAGAATA--------------
PMONBT        .C..T.....G.....G..T..C.....T.....C.....T..G..--------------
bssyn         .C..C..C..G..C..G..C..C..G..G.....C.....C..G..CGACCTGGAGAGGG

BTHKURHD      ------------------------------------------------------------
PMONBT        ------------------------------------------------------------
bssyn         CTCAGAAGGCCGTGAACGAGCTGTTCACCAGCAGCAACCAGATCGGCCTGAAGACCGACG

BTHKURHD      ----------------------------T
PMONBT        ----------------------------C
bssyn         TGACCGACTACCACATCGATCAGGTGTAG
```

Fig. 5A

```
                   10        20        30        40        50        60
                    *         *         *         *         *         *
PMONBT   ATGGACAACAACCCAAACATCAACGAATGCATTCCATACAACTGCTTGAGTAACCCAGAA
bssyn    ..............C..........G......C..C.........C....C.....C..G

                   70        80        90       100       110       120
                    *         *         *         *         *         *
PMONBT   GTTGAAGTACTTGGTGGAGAACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG
bssyn    ..G..G..G..G..C..C..G.....C..G.....C.....C............AG.C..

                  130       140       150       160       170       180
                    *         *         *         *         *         *
PMONBT   TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTGCTGGGTTCGTTCTCGGACTA
bssyn    AG.C....C.....C.....G..............C..C..C..C.....G..G..C..G

                  190       200       210       220       230       240
                    *         *         *         *         *         *
PMONBT   GTTGACATCATCTGGGTATCTTTGGTCCATCTCAATGGGATGCATTCCTGGTGCAAATT
bssyn    ..G..............C.....C..C..CAGC..G.....C..C...........G..C

                  250       260       270       280       290       300
                    *         *         *         *         *         *
PMONBT   GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGAACCAGGCCATCTCTAGGTTG
bssyn    ......C...........C.C.....G.........C.C............AGCC.CC..

                  310       320       330       340       350       360
                    *         *         *         *         *         *
PMONBT   GAAGGATTGAGCAATCTCTACCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT
bssyn    ..G..CC.......C..G...........C..C.........C.C........G.....C

                  370       380       390       400       410       420
                    *         *         *         *         *         *
PMONBT   CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAATTCAACGACATGAACAGCGCC
bssyn    ..C..C.....C..C..G........G.....C..C..G.....................

                  430       440       450       460       470       480
                    *         *         *         *         *         *
PMONBT   TTGACCACAGCTATCCCATTGTTCGCAGTCCAGAACTACCAAGTTCCTCTCTTGTCCGTG
bssyn    C.......C..C......CC......C..G...........G..G..C..GC..AG....

                  490       500       510       520       530       540
                    *         *         *         *         *         *
PMONBT   TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAGACGTTAGCGTGTTTGGGCAA
bssyn    .....G..G..C..C..C..G.....G........G..C.....C........C..C..G

                  550       560       570       580       590       600
                    *         *         *         *         *         *
PMONBT   AGGTGGGGATTCGATGCTGCAACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT
bssyn    C.C.....C.....C..C..C........C.....C...........G..CC.C.....C

                  610       620       630       640       650       660
                    *         *         *         *         *         *
PMONBT   GGAAACTACACCGACCACGCTGTTCGTTGGTACAACACTGGCTTGGAGCGTGTCTGGGGT
bssyn    ..C.................C..G..C...........C...C.......C..G......

                  670       680       690       700       710       720
                    *         *         *         *         *         *
PMONBT   CCTGATTCTAGAGATTGGATTAGATACAACCAGTTCAGGAGAGAATTGACCCTCACAGTT
bssyn    ..C..CAGCC.C..C.....C..G............C.CC.C..GC.......G..C..G
```

Fig. 5B

```
                 730       740       750       760       770       780
                   *         *         *         *         *         *
PMONBT  TTGGACATTGTGTCTCTCTTCCCGAACTATGACTCCAGAACCTACCCTATCCGTACAGTG
bssyn   C.......C...AGC..G.....C.....C...AG.C.C........C.....C..C...

                 790       800       810       820       830       840
                   *         *         *         *         *         *
PMONBT  TCCCAACTTACCAGAGAAATCTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC
bssyn   AG...G..G...C.C..G..T..C..C.....C..G..G..............C......

                 850       860       870       880       890       900
                   *         *         *         *         *         *
PMONBT  CGTGGTTCTGCCCAAGGTATCGAAGGCTCCATCAGGAGCCCACACTTGATGGACATCTTG
bssyn   ..C..CAGC.....G..C.....G...AG....C.C.....C...C...........C..

                 910       920       930       940       950       960
                   *         *         *         *         *         *
PMONBT  AACAGCATAACTATCTACAGCGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG
bssyn   ........C..C.......C...C..C...C.C..C.....C......AGC..C......

                 970       980       990      1000      1010      1020
                   *         *         *         *         *         *
PMONBT  ATCATGGCCTCTCCAGTTGGATTCAGCGGGCCCGAGTTTACCTTTCCTCTCTATGGAACT
bssyn   .........AGC..C..C..C........C........C.....C..C..G..C..C..C

                1030      1040      1050      1060      1070      1080
                   *         *         *         *         *         *
PMONBT  ATGGGAAACGCCGCTCCACAACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA
bssyn   .....C.....T..A..T..G..G..C.....G..A..G..G..C.....A..G...C.C

                1090      1100      1110      1120      1130      1140
                   *         *         *         *         *         *
PMONBT  ACCTTGTCTTCCACCTTGTACAGAAGACCCTTCAATATCGGTATCAACAACCAGCAACTT
bssyn   ...C..AGCAG....C.....C.TC....T.....C.....C..............G..G

                1150      1160      1170      1180      1190      1200
                   *         *         *         *         *         *
PMONBT  TCCGTTCTTGACGGAACAGAGTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT
bssyn   AG...G..G.....C..C...........C..C...AGCAGC...C....CAG...C..G

                1210      1220      1230      1240      1250      1260
                   *         *         *         *         *         *
PMONBT  TACAGAAAGAGCGGAACCGTTGATTCCTTGGACGAAATCCCACCACAGAACAACAATGTG
bssyn   ...C.C........C.....G..CAG.C.......G.....C..T...........C...

                1270      1280      1290      1300      1310      1320
                   *         *         *         *         *         *
PMONBT  CCACCCAGGCAAGGATTCTCCCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATTC
bssyn   .....TC.A..G..C...AG....C.TC...........AG.........CAGT..C...

                1330      1340      1350      1360      1370      1380
                   *         *         *         *         *         *
PMONBT  AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCTCATGGATTCATCGTAGTGCT
bssyn   ........CAG.............C.T..A.........AGC........C..C.....C

                1390      1400      1410      1420      1430      1440
                   *         *         *         *         *         *
PMONBT  GAGTTCAACAATATCATTCCTTCCTCTCAAATCACCCAAATCCCATTGACCAAGTCTACT
bssyn   ...........C.....C..CAG.AGC..G........G.....CC........AGC..C
```

Fig. 5C

```
                1450      1460      1470      1480      1490      1500
                   *         *         *         *         *         *
PMONBT    AACCTTGGATCTGGAACTTCTGTCGTGAAAGGACCAGGCTTCACAGGAGGTGATATTCTT
bssyn     .....G..CAGC..C..CAGC..G.....G..C..C........C..C..C..C..C..G

                1510      1520      1530      1540      1550      1560
                   *         *         *         *         *         *
PMONBT    AGAAGAACTTCTCCTGGCCAGATTAGCACCCTCAGAGTTAACATCACTGCACCACTTTCT
bssyn     C.CC.C..CAGC..C........C........GC.C..G........C..C..C..GAGC

                1570      1580      1590      1600      1610      1620
                   *         *         *         *         *         *
PMONBT    CAAAGATATCGTGTCAGGATTCGTTACGCATCTACCACTAACTTGCAATTCCACACCTCC
bssyn     ..GC.C..C..C...C.C..C..C.....CAGC.....C...C....G.........AG.

                1630      1640      1650      1660      1670      1680
                   *         *         *         *         *         *
PMONBT    ATCGACGGAAGGCCTATCAATCAGGGTAACTTCTCCGCAACCATGTCAAGCGGCAGCAAC
bssyn     ........CC.C..C.....C.....C......AG...C......AGC............

                1690      1700      1710      1720      1730      1740
                   *         *         *         *         *         *
PMONBT    TTGCAATCCGGCAGCTTCAGAACCGTCGGTTTCACTACTCCTTTCAACTTCTCTAACGGA
bssyn     C....GAG..........C.C.....G..C.....C..C..C.........AGC.....C

                1750      1760      1770      1780      1790      1800
                   *         *         *         *         *         *
PMONBT    TCAAGCGTTTTCACCCTTAGCGCTCATGTGTTCAATTCTGGCAATGAAGTGTACATTGAC
bssyn     AGC.....G........G.....C..C........CAGC.....C..G........C...

                1810      1820      1830      1840
                   *         *         *         *
PMONBT    CGTATTGAGTTTGTGCCTGCCGAAGTTACCTTCGAGGCTGAGTA----------------
bssyn     ..C..C.....C.....C.....G..G..........C.....CGACCTGGAGAGGGCT

PMONBT    ------------------------------------------------------------
bssyn     CAGAAGGCCGTGAACGAGCTGTTCACCAGCAGCAACCAGATCGGCCTGAAGACCGACGTG

PMONBT    --------------------------C
bssyn     ACCGACTACCACATCGATCAGGTGTAG
```

Fig. 6A

```
  64  ATGGACCTGC TGCCCGACGC CCGCATCGAG GACAGCCTGT GCATCGCCGA GGGCAACAAC
      MetAspLeu LeuProAsp AlaArgIleGlu AspSerLeu CysIleAla GluGlyAsnAsn

 124  ATCGACCCCT TCGTGAGCGC CAGCACCGTG CAGACCGGCA TCAACATCGC CGGCCGCATC
      IleAspPro PheValSer AlaSerThrVal GlnThrGly IleAsnIle AlaGlyArgIle

 184  CTGGGCGTGC TGGGCGTGCC CTTCGCCGGC CAGCTGGCCA GCTTCTACAG CTTCCTGGTG
      LeuGlyVal LeuGlyVal ProPheAlaGly GlnLeuAla SerPheTyr SerPheLeuVal

 244  GGCGAGCTGT GGCCCCGCGG CCGCGACCAG TGGGAGATCT TCCTGGAGCA CGTGGAGCAG
      GlyGluLeu TrpProArg GlyArgAspGln TrpGluIle PheLeuGlu HisValGluGln

 304  CTGATCAACC AGCAGATCAC CGAGAACGCC CGCAACACCG CCCTGGCCCG CCTGCAGGGC
      LeuIleAsn GlnGlnIle ThrGluAsnAla ArgAsnThr AlaLeuAla ArgLeuGlnGly

 364  CTGGGCGACA GCTTCCGCGC CTACCAGCAG AGCCTGGAGG ACTGGCTGGA GAACCGCGAC
      LeuGlyAsp SerPheArg AlaTyrGlnGln SerLeuGlu AspTrpLeu GluAsnArgAsp

 424  GACGCCCGCA CCCGCAGCGT GCTGTACACC CAGTACATCG CCCTGGAGCT GGACTTCCTG
      AspAlaArg ThrArgSer ValLeuTyrThr GlnTyrIle AlaLeuGlu LeuAspPheLeu

 484  AACGCCATGC CCCTGTTCGC CATCCGCAAC CAGGAGGTGC CCCTGCTGAT GGTGTACGCC
      AsnAlaMet ProLeuPhe AlaIleArgAsn GlnGluVal ProLeuLeu MetValTyrAla

 544  CAGGCCGCCA ACCTGCACCT GCTGCTGCTG CGCGACGCCA GCCTGTTCGG CAGCGAGTTC
      GlnAlaAla AsnLeuHis LeuLeuLeuLeu ArgAspAla SerLeuPhe GlySerGluPhe

 604  GGCCTGACCA GCCAGGAGAT CCAGCGCTAC TACGAGCGCC AGGTGGAGCG CACCCGCGAC
      GlyLeuThr SerGlnGlu IleGlnArgTyr TyrGluArg GlnValGlu ArgThrArgAsp

 664  TACAGCGACT ACTGCGTGGA GTGGTACAAC ACCGGCCTGA ACAGCCTGCG CGGCACCAAC
      TyrSerAsp TyrCysVal GluTrpTyrAsn ThrGlyLeu AsnSerLeu ArgGlyThrAsn

 724  GCCGCCAGCT GGGTGCGCTA CAACCAGTTC CGCCGCGACC TGACCCTGGG CGTGCTGGAC
      AlaAlaSer TrpValArg TyrAsnGlnPhe ArgArgAsp LeuThrLeu GlyValLeuAsp

 784  CTGGTGGCCC TGTTCCCCAG CTACGACACC CGCACCTACC CCATCAACAC CAGCGCCCAG
      LeuValAla LeuPhePro SerTyrAspThr ArgThrTyr ProIleAsn ThrSerAlaGln

 844  CTGACCCGCG AGGTGTACAC CGACGCCATC GGCGCCACCG GCGTGAACAT GGCCAGCATG
      LeuThrArg GluValTyr ThrAspAlaIle GlyAlaThr GlyValAsn MetAlaSerMet

 904  AACTGGTACA ACAACAACGC CCCCAGCTTC AGCGCCATCG AGGCCGCCGC CATCCGCAGC
      AsnTrpTyr AsnAsnAsn AlaProSerPhe SerAlaIle GluAlaAla AlaIleArgSer

 964  CCCCACCTGC TGGACTTCCT GGAGCAGCTG ACCATCTTCA GCGCCAGCAG CCGCTGGAGC
      ProHisLeu LeuAspPhe LeuGluGlnLeu ThrIlePhe SerAlaSer SerArgTrpSer

1024  AACACCCGCC ACATGACCTA CTGGCGCGGC CACACCATCC AGAGCCGCCC CATCGGCGGC
      AsnThrArg HisMetThr TyrTrpArgGly HisThrIle GlnSerArg ProIleGlyGly
```

Fig. 6B

```
1084  GGCCTGAACA CCAGCACCCA CGGCGCCACC AACACCAGCA TCAACCCCGT GACCCTGCGC
      GlyLeuAsn ThrSerThr HisGlyAlaThr AsnThrSer IleAsnPro ValThrLeuArg

1144  TTCGCCAGCC GCGACGTGTA CCGCACCGAG AGCTACGCCG GCGTGCTGCT GTGGGGCATC
      PheAlaSer ArgAspVal TyrArgThrGlu SerTyrAla GlyValLeu LeuTrpGlyIle

1204  TACCTGGAGC CCATCCACGG CGTGCCCACC GTGCGCTTCA ACTTCACCAA CCCCCAGAAC
      TyrLeuGlu ProIleHis GlyValProThr ValArgPhe AsnPheThr AsnProGlnAsn

1264  ATCAGCGACC GCGGCACCGC CAACTACAGC CAGCCCTACG AGAGCCCCGG CCTGCAGCTG
      IleSerAsp ArgGlyThr AlaAsnTyrSer GlnProTyr GluSerPro GlyLeuGlnLeu

1324  AAGGACAGCG AGACCGAGCT GCCCCCCGAG ACCACCGAGC GCCCCAACTA CGAGAGCTAC
      LysAspSer GluThrGlu LeuProProGlu ThrThrGlu ArgProAsn TyrGluSerTyr

1384  AGCCACCGCC TGAGCCACAT CGGCATCATC CTGCAGAGCC GCGTGAACGT GCCCGTGTAC
      SerHisArg LeuSerHis IleGlyIleIle LeuGlnSer ArgValAsn ValProValTyr

1444  AGCTGGACCC ACCGCAGCGC CGACCGCACC AACACCATCG GCCCCAACCG CATCACCCAG
      SerTrpThr HisArgSer AlaAspArgThr AsnThrIle GlyProAsn ArgIleThrGln

1504  ATCCCCATGG TGAAGGCCAG CGAGCTGCCC CAGGGCACCA CCGTGGTGCG CGGCCCCGGC
      IleProMet ValLysAla SerGluLeuPro GlnGlyThr ThrValVal ArgGlyProGly

1564  TTCACCGGCG GCGACATCCT GCGCCGCACC AACACCGGCG GCTTCGGCCC CATCCGCGTG
      PheThrGly GlyAspIle LeuArgArgThr AsnThrGly GlyPheGly ProIleArgVal

1624  ACCGTGAACG GCCCCCTGAC CCAGCGCTAC CGCATCGGCT TCCGCTACGC CAGCACCGTG
      ThrValAsn GlyProLeu ThrGlnArgTyr ArgIleGly PheArgTyr AlaSerThrVal

1684  GACTTCGACT TCTTCGTGAG CCGCGGCGGC ACCACCGTGA ACAACTTCCG CTTCCTGCGC
      AspPheAsp PhePheVal SerArgGlyGly ThrThrVal AsnAsnPhe ArgPheLeuArg

1744  ACCATGAACA GCGGCGACGA GCTGAAGTAC GGCAACTTCG TGCGCCGCGC CTTCACCACC
      ThrMetAsn SerGlyAsp GluLeuLysTyr GlyAsnPhe ValArgArg AlaPheThrThr

1804  CCCTTCACCT TCACCCAGAT CCAGGACATC ATCCGCACCA GCATCCAGGG CCTGAGCGGC
      ProPheThr PheThrGln IleGlnAspIle IleArgThr SerIleGln GlyLeuSerGly

1864  AACGGCGAGG TGTACATCGA CAAGATCGAG ATCATCCCCG TGACCGCCAC CTTCGAGGCC
      AsnGlyGlu ValTyrIle AspLysIleGlu IleIlePro ValThrAla ThrPheGluAla

1924  GAGTACGACC TGGAGCGCGC CCAGGAGGCC GTGAACGCCC TGTTCACCAA CACCAACCCC
      GluTyrAsp LeuGluArg AlaGlnGluAla ValAsnAla LeuPheThr AsnThrAsnPro

1984  CGCCGCCTGA AGACCGACGT GACCGACTAC CACATCGACC AGGTGAGCAA CCTGGTGGCC
      ArgArgLeu LysThrAsp ValThrAspTyr HisIleAsp GlnValSer AsnLeuValAla

2044  TGCCTGAGCG ACGAGTTCTG CCTGGACGAG AAGCGCGAGC TGCTGGAGAA GGTGAAGTAC
      CysLeuSer AspGluPhe CysLeuAspGlu LysArgGlu LeuLeuGlu LysValLysTyr
```

Fig. 6C

```
2104  GCCAAGCGCC TGAGCGACGA GCGCAACCTG CTGCAGGACC CCAACTTCAC CAGCATCAAC
      AlaLysArg LeuSerAsp GluArgAsnLeu LeuGlnAsp ProAsnPhe ThrSerIleAsn

2164  AAGCAGCCCG ACTTCATCAG CACCAACGAG CAGAGCAACT TCACCAGCAT CCACGAGCAG
      LysGlnPro AspPheIle SerThrAsnGlu GlnSerAsn PheThrSer IleHisGluGln

2224  AGCGAGCACG GCTGGTGGGG CAGCGAGAAC ATCACCATCC AGGAGGGCAA CGACGTGTTC
      SerGluHis GlyTrpTrp GlySerGluAsn IleThrIle GlnGluGly AsnAspValPhe

2284  AAGGAGAACT ACGTGACCCT GCCCGGCACC TTCAACGAGT GCTACCCCAC CTACCTGTAC
      LysGluAsn TyrValThr LeuProGlyThr PheAsnGlu CysTyrPro ThrTyrLeuTyr

2344  CAGAAGATCG GCGAGAGCGA GCTGAAGGCC TACACCCGCT ACCAGCTGCG CGGCTACATC
      GlnLysIle GlyGluSer GluLeuLysAla TyrThrArg TyrGlnLeu ArgGlyTyrIle

2404  GAGGACAGCC AGGACCTGGA GATCTACCTG ATCCGCTACA ACGCCAAGCA CGAGACCCTG
      GluAspSer GlnAspLeu GluIleTyrLeu IleArgTyr AsnAlaLys HisGluThrLeu

2464  GACGTGCCCG GCACCGAGAG CCTGTGGCCC CTGAGCGTGG AGAGCCCCAT CGGCCGCTGC
      AspValPro GlyThrGlu SerLeuTrpPro LeuSerVal GluSerPro IleGlyArgCys

2524  GGCGAGCCCA ACCGCTGCGC CCCCCACTTC GAGTGGAACC CCGACCTGGA CTGCAGCTGC
      GlyGluPro AsnArgCys AlaProHisPhe GluTrpAsn ProAspLeu AspCysSerCys

2584  CGCGACGGCG AGAAGTGCGC CCACCACAGC CACCACTTCA GCCTGGACAT CGACGTGGGC
      ArgAspGly GluLysCys AlaHisHisSer HisHisPhe SerLeuAsp IleAspValGly

2644  TGCACCGACC TGCACGAGAA CCTGGGCGTG TGGGTGGTGT TCAAGATCAA GACCCAGGAG
      CysThrAsp LeuHisGlu AsnLeuGlyVal TrpValVal PheLysIle LysThrGlnGlu

2704  GGCCACGCCC GCCTGGGCAA CCTGGAGTTC ATCGAGGAGA AGCCCCTGCT GGGCGAGGCC
      GlyHisAla ArgLeuGly AsnLeuGluPhe IleGluGlu LysProLeu LeuGlyGluAla

2764  CTGAGCCGCG TGAAGCGCGC CGAGAAGAAG TGGCGCGACA AGCGCGAGAA GCTGCAGCTG
      LeuSerArg ValLysArg AlaGluLysLys TrpArgAsp LysArgGlu LysLeuGlnLeu

2824  GAGACCAAGC GCGTGTACAC CGAGGCCAAG GAGGCCGTGG ACGCCCTGTT CGTGGACAGC
      GluThrLys ArgValTyr ThrGluAlaLys GluAlaVal AspAlaLeu PheValAspSer

2884  CAGTACGACC GCCTGCAGGC CGACACCAAC ATCGGCATGA TCCACGCCGC CGACAAGCTG
      GlnTyrAsp ArgLeuGln AlaAspThrAsn IleGlyMet IleHisAla AlaAspLysLeu

2944  GTGCACCGCA TCCGCGAGGC CTACCTGAGC GAGCTGCCCG TGATCCCCGG CGTGAACGCC
      ValHisArg IleArgGlu AlaTyrLeuSer GluLeuPro ValIlePro GlyValAsnAla

3004  GAGATCTTCG AGGAGCTGGA GGGCCACATC ATCACCGCCA TCAGCCTGTA CGACGCCCGC
      GluIlePhe GluGluLeu GluGlyHisIle IleThrAla IleSerLeu TyrAspAlaArg
```

Fig. 6D

```
3064  AACGTGGTGA AGAACGGCGA CTTCAACAAC GGCCTGACCT GCTGGAACGT GAAGGGCCAC
      AsnValVal LysAsnGly AspPheAsnAsn GlyLeuThr CysTrpAsn ValLysGlyHis

3124  GTGGACGTGC AGCAGAGCCA CCACCGCAGC GACCTGGTGA TCCCCGAGTG GGAGGCCGAG
      ValAspVal GlnGlnSer HisHisArgSer AspLeuVal IleProGlu TrpGluAlaGlu

3184  GTGAGCCAGG CCGTGCGCGT GTGCCCCGGC TGCGGCTACA TCCTGCGCGT GACCGCCTAC
      ValSerGln AlaValArg ValCysProGly CysGlyTyr IleLeuArg ValThrAlaTyr

3244  AAGGAGGGCT ACGGCGAGGG CTGCGTGACC ATCCACGAGA TCGAGAACAA CACCGACGAG
      LysGluGly TyrGlyGlu GlyCysValThr IleHisGlu IleGluAsn AsnThrAspGlu

3304  CTGAAGTTCA AGAACCGCGA GGAGGAGGAG GTGTACCCCA CCGACACCGG CACCTGCAAC
      LeuLysPhe LysAsnArg GluGluGluGlu ValTyrPro ThrAspThr GlyThrCysAsn

3364  GACTACACCG CCCACCAGGG CACCGCCGGC TGCGCCGACG CCTGCAACAG CCGCAACGCC
      AspTyrThr AlaHisGln GlyThrAlaGly CysAlaAsp AlaCysAsn SerArgAsnAla

3424  GGCTACGAGG ACGCCTACGA GGTGGACACC ACCGCCAGCG TGAACTACAA GCCCACCTAC
      GlyTyrGlu AspAlaTyr GluValAspThr ThrAlaSer ValAsnTyr LysProThrTyr

3484  GAGGAGGAGA CCTACACCGA CGTGCGCCGC GACAACCACT GCGAGTACGA CCGCGGCTAC
      GluGluGlu ThrTyrThr AspValArgArg AspAsnHis CysGluTyr AspArgGlyTyr

3544  GTGAACTACC CCCCCGTGCC CGCCGGCTAC GTGACCAAGG AGCTGGAGTA CTTCCCCGAG
      ValAsnTyr ProProVal ProAlaGlyTyr ValThrLys GluLeuGlu TyrPheProGlu

3604  ACCGACACCG TGTGGATCGA GATCGGCGAG ACCGAGGGCA AGTTCATCGT GGACAGCGTG
      ThrAspThr ValTrpIle GluIleGlyGlu ThrGluGly LysPheIle ValAspSerVal

3664  GAGCTGCTGC TGATGGAGGA GTAG
      GluLeuLeu LeuMetGlu Glu---
```

Fig. 7A

SEQUENCE OF THE FULL-LENGTH HYBRID SYNTHETIC/NATIVE CRYIA(B) CHIMERIC GENE
The fusion point between the synthetic and native coding sequences is indicated by a slash (/) in the sequence.

```
  1  ATGGACAACA ACCCCAACAT CAACGAGTGC ATCCCCTACA ACTGCCTGAG CAACCCCGAG
     MetAspAsn AsnProAsn IleAsnGluCys IleProTyr AsnCysLeu SerAsnProGlu

 61  GTGGAGGTGC TGGGCGGCGA GCGCATCGAG ACCGGCTACA CCCCCATCGA CATCAGCCTG
     ValGluVal LeuGlyGly GluArgIleGlu ThrGlyTyr ThrProIle AspIleSerLeu

121  AGCCTGACCC AGTTCCTGCT GAGCGAGTTC GTGCCCGGCG CCGGCTTCGT GCTGGGCCTG
     SerLeuThr GlnPheLeu LeuSerGluPhe ValProGly AlaGlyPhe ValLeuGlyLeu

181  GTGGACATCA TCTGGGGCAT CTTCGGCCCC AGCCAGTGGG ACGCCTTCCT GGTGCAGATC
     ValAspIle IleTrpGly IlePheGlyPro SerGlnTrp AspAlaPhe LeuValGlnIle

241  GAGCAGCTGA TCAACCAGCG CATCGAGGAG TTCGCCCGCA ACCAGGCCAT CAGCCGCCTG
     GluGlnLeu IleAsnGln ArgIleGluGlu PheAlaArg AsnGlnAla IleSerArgLeu

301  GAGGGCCTGA GCAACCTGTA CCAAATCTAC GCCGAGAGCT TCCGCGAGTG GGAGGCCGAC
     GluGlyLeu SerAsnLeu TyrGlnIleTyr AlaGluSer PheArgGlu TrpGluAlaAsp

361  CCCACCAACC CCGCCCTGCG CGAGGAGATG CGCATCCAGT TCAACGACAT GAACAGCGCC
     ProThrAsn ProAlaLeu ArgGluGluMet ArgIleGln PheAsnAsp MetAsnSerAla

421  CTGACCACCG CCATCCCCCT GTTCGCCGTG CAGAACTACC AGGTGCCCCT GCTGAGCGTG
     LeuThrThr AlaIlePro LeuPheAlaVal GlnAsnTyr GlnValPro LeuLeuSerVal

481  TACGTGCAGG CCGCCAACCT GCACCTGAGC GTGCTGCGCG ACGTCAGCGT GTTCGGCCAG
     TyrValGln AlaAlaAsn LeuHisLeuSer ValLeuArg AspValSer ValPheGlyGln

541  CGCTGGGGCT TCGACGCCGC CACCATCAAC AGCCGCTACA ACGACCTGAC CCGCCTGATC
     ArgTrpGly PheAspAla AlaThrIleAsn SerArgTyr AsnAspLeu ThrArgLeuIle

601  GGCAACTACA CCGACCACGC CGTGCGCTGG TACAACACCG GCCTGGAGCG CGTGTGGGGT
     GlyAsnTyr ThrAspHis AlaValArgTrp TyrAsnThr GlyLeuGlu ArgValTrpGly

661  CCCGACAGCC GCGACTGGAT CAGGTACAAC CAGTTCCGCC GCGAGCTGAC CCTGACCGTG
     ProAspSer ArgAspTrp IleArgTyrAsn GlnPheArg ArgGluLeu ThrLeuThrVal

721  CTGGACATCG TGAGCCTGTT CCCCAACTAC GACAGCCGCA CCTACCCCAT CCGCACCGTG
     LeuAspIle ValSerLeu PheProAsnTyr AspSerArg ThrTyrPro IleArgThrVal

781  AGCCAGCTGA CCCGCGAGAT TTACACCAAC CCCGTGCTGG AGAACTTCGA CGGCAGCTTC
     SerGlnLeu ThrArgGlu IleTyrThrAsn ProValLeu GluAsnPhe AspGlySerPhe

841  CGCGGCAGCG CCCAGGGCAT CGAGGGCAGC ATCCGCAGCC CCCACCTGAT GGACATCCTG
     ArgGlySer AlaGlnGly IleGluGlySer IleArgSer ProHisLeu MetAspIleLeu

901  AACAGCATCA CCATCTACAC CGACGCCCAC CGCGGCGAGT ACTACTGGAG CGGCCACCAG
     AsnSerIle ThrIleTyr ThrAspAlaHis ArgGlyGlu TyrTyrTrp SerGlyHisGln
```

Fig. 7B

```
961   ATCATGGCCA GCCCCGTCGG CTTCAGCGGC CCCGAGTTCA CCTTCCCCCT GTACGGCACC
      IleMetAla SerProVal GlyPheSerGly ProGluPhe ThrPhePro LeuTyrGlyThr

1021  ATGGGCAACG CTGCACCTCA GCAGCGCATC GTGGCACAGC TGGGCCAGGG AGTGTACCGC
      MetGlyAsn AlaAlaPro GlnGlnArgIle ValAlaGln LeuGlyGln GlyValTyrArg

1081  ACCCTGAGCA GCACCCTGTA CCGTCGACCT TTCAACATCG GCATCAACAA CCAGCAGCTG
      ThrLeuSer SerThrLeu TyrArgArgPro PheAsnIle GlyIleAsn AsnGlnGlnLeu

1141  AGCGTGCTGG ACGGCACCGA GTTCGCCTAC GGCACCAGCA GCAACCTGCC CAGCGCCGTG
      SerValLeu AspGlyThr GluPheAlaTyr GlyThrSer SerAsnLeu ProSerAlaVal

1201  TACCGCAAGA GCGGCACCGT GGACAGCCTG GACGAGATCC CCCCTCAGAA CAACAACGTG
      TyrArgLys SerGlyThr ValAspSerLeu AspGluIle ProProGln AsnAsnAsnVal

1261  CCACCTCGAC AGGGCTTCAG CCACCGTCTG AGCCACGTGA GCATGTTCCG CAGTGGCTTC
      ProProArg GlnGlyPhe SerHisArgLeu SerHisVal SerMetPhe ArgSerGlyPhe

1321  AGCAACAGCA GCGTGAGCAT CATCCGTGCA CCTATGTTCA GCTGGATTCA CCGCAGTGCC
      SerAsnSer SerValSer IleIleArgAla ProMetPhe SerTrpIle HisArgSerAla

1381  GAGTTCAACA ACATCATCCC CAGCAGCCAG ATCACCCAGA TCCCCCTGAC CAAGAGCACC
      GluPheAsn AsnIleIle ProSerSerGln IleThrGln IleProLeu ThrLysSerThr

1441  AACCTGGGCA GCGGCACCAG CGTGGTGAAG GGCCCCGGCT TCACCGGCGG CGACATCCTG
      AsnLeuGly SerGlyThr SerValValLys GlyProGly PheThrGly GlyAspIleLeu

1501  CGCCGCACCA GCCCCGGCCA GATCAGCACC CTGCGCGTGA ACATCACCGC CCCCCTGAGC
      ArgArgThr SerProGly GlnIleSerThr LeuArgVal AsnIleThr AlaProLeuSer

1561  CAGCGCTACC GCGTCCGCAT CCGCTACGCC AGCACCACCA ACCTGCAGTT CCACACCAGC
      GlnArgTyr ArgValArg IleArgTyrAla SerThrThr AsnLeuGln PheHisThrSer

1621  ATCGACGGCC GCCCCATCAA CCAGGGCAAC TTCAGCGCCA CCATGAGCAG CGGCAGCAAC
      IleAspGly ArgProIle AsnGlnGlyAsn PheSerAla ThrMetSer SerGlySerAsn

1681  CTGCAGAGCG GCAGCTTCCG CACCGTGGGC TTCACCACCC CCTTCAACTT CAGCAACGGC
      LeuGlnSer GlySerPhe ArgThrValGly PheThrThr ProPheAsn PheSerAsnGly

1741  AGCAGCGTGT TCACCCTGAG CGCCCACGTG TTCAACAGCG GCAACGAGGT GTACATCGAC
      SerSerVal PheThrLeu SerAlaHisVal PheAsnSer GlyAsnGlu ValTyrIleAsp

1801  CGCATCGAGT TCGTGCCCGC CGAGGTGACC TTCGAGGCCG AGTACGACCT GGAGAGGGCT
      ArgIleGlu PheValPro AlaGluValThr PheGluAla GluTyrAsp LeuGluArgAla

1861  CAGAAGGCCG TGAACGAGCT GTTCACCAGC AGCAACCAGA TCGGCCTGAA GACCGACGTG
      GlnLysAla ValAsnGlu LeuPheThrSer SerAsnGln IleGlyLeu LysThrAspVal
```

Fig. 7C

```
1921  ACCGACTACC ACATCGAT/CA AGTATCCAAT TTAGTTGAGT GTTTATCTGATGAATTTTGT
      ThrAspTyr HisIleAsp/GlnValSerAsn LeuValGlu CysLeuSer AspGluPheCys

1981  CTGGATGAAA AAAAAGAATT GTCCGAGAAA GTCAAACATG CGAAGCGACT TAGTGATGAG
      LeuAspGlu LysLysGlu LeuSerGluLys ValLysHis AlaLysArg LeuSerAspGlu

2041  CGGAATTTAC TTCAAGATCC AAACTTTAGA GGGATCAATA GACAACTAGA CCGTGGCTGG
      ArgAsnLeu LeuGlnAsp ProAsnPheArg GlyIleAsn ArgGlnLeu AspArgGlyTrp

2101  AGAGGAAGTA CGGATATTAC CATCCAAGGA GGCGATGACG TATTCAAAGA GAATTACGTT
      ArgGlySer ThrAspIle ThrIleGlnGly GlyAspAsp ValPheLys GluAsnTyrVal

2161  ACGCTATTGG GTACCTTTGA TGAGTGCTAT CCAACGTATT TATATCAAAA AATAGATGAG
      ThrLeuLeu GlyThrPhe AspGluCysTyr ProThrTyr LeuTyrGln LysIleAspGlu

2221  TCGAAATTAA AAGCCTATAC CCGTTACCAA TTAAGAGGGT ATATCGAAGA TAGTCAAGAC
      SerLysLeu LysAlaTyr ThrArgTyrGln LeuArgGly TyrIleGlu AspSerGlnAsp

2281  TTAGAAATCT ATTTAATTCG CTACAATGCC AAACACGAAA CAGTAAATGT GCCAGGTACG
      LeuGluIle TyrLeuIle ArgTyrAsnAla LysHisGlu ThrValAsn ValProGlyThr

2341  GGTTCCTTAT GGCCGCTTTC AGCCCCAAGT CCAATCGGAA AATGTGCCCA TCATTCCCAT
      GlySerLeu TrpProLeu SerAlaProSer ProIleGly LysCysAla HisHisSerHis

2401  CATTTCTCCT TGGACATTGA TGTTGGATGT ACAGACTTAA ATGAGGACTT AGGTGTATGG
      HisPheSer LeuAspIle AspValGlyCys ThrAspLeu AsnGluAsp LeuGlyValTrp

2461  GTGATATTCA AGATTAAGAC GCAAGATGGC CATGCAAGAC TAGGAAATCT AGAATTTCTC
      ValIlePhe LysIleLys ThrGlnAspGly HisAlaArg LeuGlyAsn LeuGluPheLeu

2521  GAAGAGAAAC CATTAGTAGG AGAAGCACTA GCTCGTGTGA AAAGAGCGGA GAAAAAATGG
      GluGluLys ProLeuVal GlyGluAlaLeu AlaArgVal LysArgAla GluLysLysTrp

2581  AGAGACAAAC GTGAAAAATT GGAATGGGAA ACAAATATTG TTTATAAAGA GGCAAAAGAA
      ArgAspLys ArgGluLys LeuGluTrpGlu ThrAsnIle ValTyrLys GluAlaLysGlu

2641  TCTGTAGATG CTTTATTTGT AAACTCTCAA TATGATAGAT TACAAGCGGA TACCAACATC
      SerValAsp AlaLeuPhe ValAsnSerGln TyrAspArg LeuGlnAla AspThrAsnIle

2701  GCGATGATTC ATGCGGCAGA TAAACGCGTT CATAGCATTC GAGAAGCTTA TCTGCCTGAG
      Ala e Ile HisAlaAla AspLysArgVal HisSerIle ArgGluAla TyrLeuProGlu

2761  CTGTCTGTGA TTCCGGGTGT CAATGCGGCT ATTTTTGAAG AATTAGAAGG GCGTATTTTC
      LeuSerVal IleProGly ValAsnAlaAla IlePheGlu GluLeuGlu GlyArgIlePhe

2821  ACTGCATTCT CCCTATATGA TGCGAGAAAT GTCATTAAAA ATGGTGATTT TAATAATGGC
      ThrAlaPhe SerLeuTyr AspAlaArgAsn ValIleLys AsnGlyAsp PheAsnAsnGly
```

Fig. 7D

```
2881  TTATCCTGCT GGAACGTGAA AGGGCATGTA GATGTAGAAG AACAAAACAA CCACCGTTCG
      LeuSerCys TrpAsnVal LysGlyHisVal AspValGlu GluGlnAsn AsnHisArgSer

2941  GTCCTTGTTG TTCCGGAATG GGAAGCAGAA GTGTCACAAG AAGTTCGTGT CTGTCCGGGT
      ValLeuVal ValProGlu TrpGluAlaGlu ValSerGln GluValArg ValCysProGly

3001  CGTGGCTATA TCCTTCGTGT CACAGCGTAC AAGGAGGGAT ATGGAGAAGG TTGCGTAACC
      ArgGlyTyr IleLeuArg ValThrAlaTyr LysGluGly TyrGlyGlu GlyCysValThr

3061  ATTCATGAGA TCGAGAACAA TACAGACGAA CTGAAGTTTA GCAACTGTGT AGAAGAGGAA
      IleHisGlu IleGluAsn AsnThrAspGlu LeuLysPhe SerAsnCys ValGluGluGlu

3121  GTATATCCAA ACAACACGGT AACGTGTAAT GATTATACTG CGACTCAAGA AGAATATGAG
      ValTyrPro AsnAsnThr ValThrCysAsn AspTyrThr AlaThrGln GluGluTyrGlu

3181  GGTACGTACA CTTCTCGTAA TCGAGGATAT GACGGAGCCT ATGAAAGCAA TTCTTCTGTA
      GlyThrTyr ThrSerArg AsnArgGlyTyr AspGlyAla TyrGluSer AsnSerSerVal

3241  CCAGCTGATT ATGCATCAGC CTATGAAGAA AAAGCATATA CAGATGGACG AAGAGACAAT
      ProAlaAsp TyrAlaSer AlaTyrGluGlu LysAlaTyr ThrAspGly ArgArgAspAsn

3301  CCTTGTGAAT CTAACAGAGG ATATGGGGAT TACACACCAC TACCAGCTGG CTATGTGACA
      ProCysGlu SerAsnArg GlyTyrGlyAsp TyrThrPro LeuProAla GlyTyrValThr

3361  AAAGAATTAG AGTACTTCCC AGAAACCGAT AAGGTATGGA TTGAGATCGG AGAAACGGAA
      LysGluLeu GluTyrPhe ProGluThrAsp LysValTrp IleGluIle GlyGluThrGlu

3421  GGAACATTCA TCGTGGACAG CGTGGAATTA CTTCTTATGG AGGAATAA
      GlyThrPhe IleValAsp SerValGluLeu LeuLeuMet GluGlu---
```

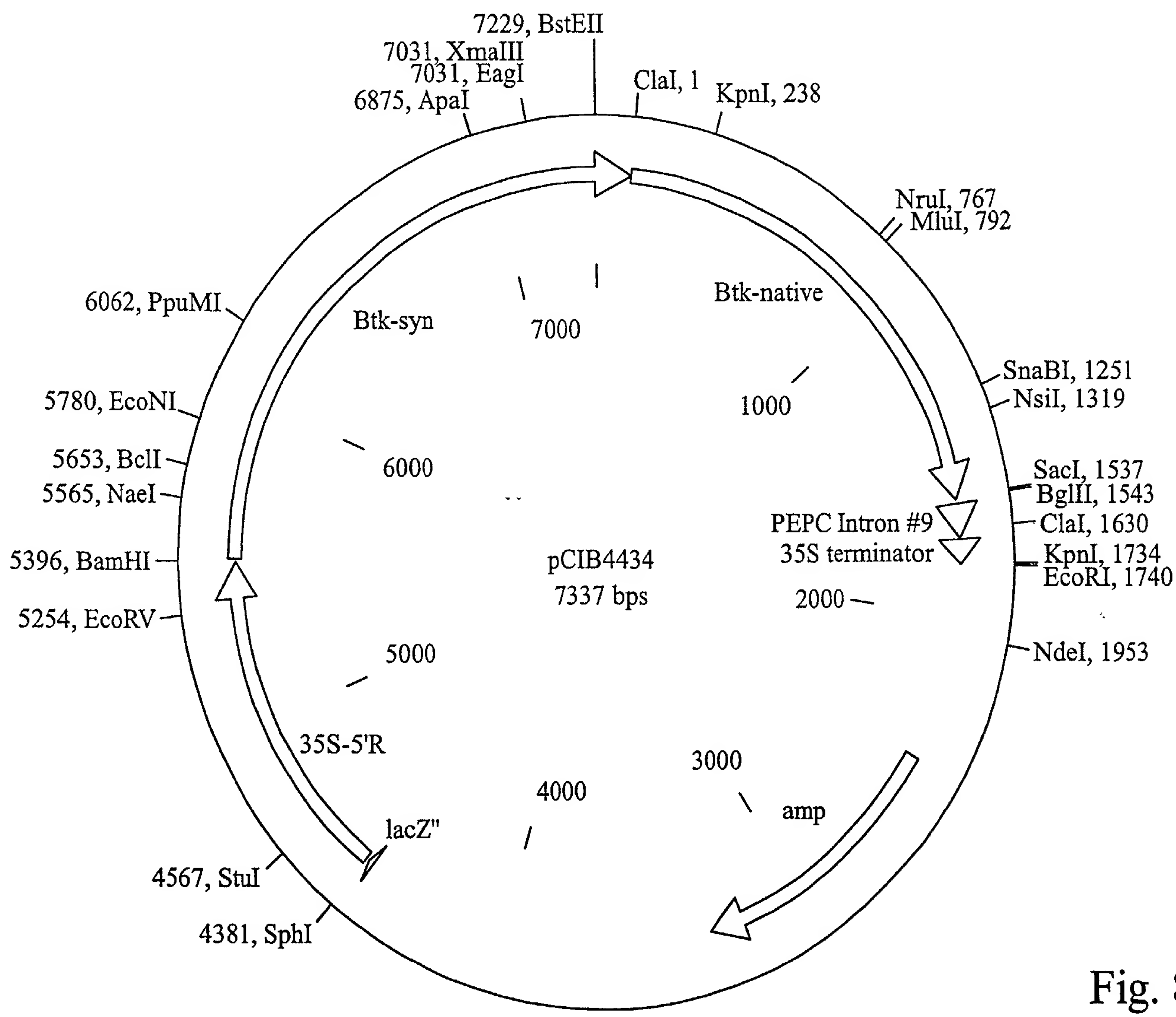

7229, BstEII
7031, XmaIII
7031, EagI
6875, ApaI
ClaI, 1
KpnI, 238
NruI, 767
MluI, 792
SnaBI, 1251
NsiI, 1319
SacI, 1537
BglII, 1543
ClaI, 1630
KpnI, 1734
EcoRI, 1740
NdeI, 1953
6062, PpuMI
5780, EcoNI
5653, BclI
5565, NaeI
5396, BamHI
5254, EcoRV
4567, StuI
4381, SphI
Btk-native
Btk-syn
PEPC Intron #9
35S terminator
amp
35S-5'R
lacZ"
pCIB4434
7337 bps
1000
2000
3000
4000
5000
6000
7000

Fig. 8

Fig. 9A

```
  1 ATGGACAACA ACCCCAACAT CAACGAGTGC ATCCCCTACA ACTGCCTGAG CAACCCCGAG
    MetAspAsn AsnProAsn IleAsnGluCys IleProTyr AsnCysLeu SerAsnProGlu

 61 GTGGAGGTGC TGGGCGGCGA GCGCATCGAG ACCGGCTACA CCCCCATCGA CATCAGCCTG
    ValGluVal LeuGlyGly GluArgIleGlu ThrGlyTyr ThrProIle AspIleSerLeu

121 AGCCTGACCC AGTTCCTGCT GAGCGAGTTC GTGCCCGGCG CCGGCTTCGT GCTGGGCCTG
    SerLeuThr GlnPheLeu LeuSerGluPhe ValProGly AlaGlyPhe ValLeuGlyLeu

181 GTGGACATCA TCTGGGGCAT CTTCGGCCCC AGCCAGTGGG ACGCCTTCCT GGTGCAGATC
    ValAspIle IleTrpGly IlePheGlyPro SerGlnTrp AspAlaPhe LeuValGlnIle

241 GAGCAGCTGA TCAACCAGCG CATCGAGGAG TTCGCCCGCA ACCAGGCCAT CAGCCGCCTG
    GluGlnLeu IleAsnGln ArgIleGluGlu PheAlaArg AsnGlnAla IleSerArgLeu

301 GAGGGCCTGA GCAACCTGTA CCAAATCTAC GCCGAGAGCT TCCGCGAGTG GGAGGCCGAC
    GluGlyLeu SerAsnLeu TyrGlnIleTyr AlaGluSer PheArgGlu TrpGluAlaAsp

361 CCCACCAACC CCGCCCTGCG CGAGGAGATG CGCATCCAGT TCAACGACAT GAACAGCGCC
    ProThrAsn ProAlaLeu ArgGluGluMet ArgIleGln PheAsnAsp MetAsnSerAla

421 CTGACCACCG CCATCCCCCT GTTCGCCGTG CAGAACTACC AGGTGCCCCT GCTGAGCGTG
    LeuThrThr AlaIlePro LeuPheAlaVal GlnAsnTyr GlnValPro LeuLeuSerVal

481 TACGTGCAGG CCGCCAACCT GCACCTGAGC GTGCTGCGCG ACGTCAGCGT GTTCGGCCAG
    TyrValGln AlaAlaAsn LeuHisLeuSer ValLeuArg AspValSer ValPheGlyGln

541 CGCTGGGGCT TCGACGCCGC CACCATCAAC AGCCGCTACA ACGACCTGAC CCGCCTGATC
    ArgTrpGly PheAspAla AlaThrIleAsn SerArgTyr AsnAspLeu ThrArgLeuIle

601 GGCAACTACA CCGACCACGC CGTGCGCTGG TACAACACCG GCCTGGAGCG CGTGTGGGGT
    GlyAsnTyr ThrAspHis AlaValArgTrp TyrAsnThr GlyLeuGlu ArgValTrpGly

661 CCCGACAGCC GCGACTGGAT CAGGTACAAC CAGTTCCGCC GCGAGCTGAC CCTGACCGTG
    ProAspSer ArgAspTrp IleArgTyrAsn GlnPheArg ArgGluLeu ThrLeuThrVal

721 CTGGACATCG TGAGCCTGTT CCCCAACTAC GACAGCCGCA CCTACCCCAT CCGCACCGTG
    LeuAspIle ValSerLeu PheProAsnTyr AspSerArg ThrTyrPro IleArgThrVal

781 AGCCAGCTGA CCCGCGAGAT TTACACCAAC CCCGTGCTGG AGAACTTCGA CGGCAGCTTC
    SerGlnLeu ThrArgGlu IleTyrThrAsn ProValLeu GluAsnPhe AspGlySerPhe

841 CGCGGCAGCG CCCAGGGCAT CGAGGGCAGC ATCCGCAGCC CCCACCTGAT GGACATCCTG
    ArgGlySer AlaGlnGly IleGluGlySer IleArgSer ProHisLeu MetAspIleLeu

901 AACAGCATCA CCATCTACAC CGACGCCCAC CGCGGCGAGT ACTACTGGAG CGGCCACCAG
    AsnSerIle ThrIleTyr ThrAspAlaHis ArgGlyGlu TyrTyrTrp SerGlyHisGln

961 ATCATGGCCA GCCCCGTCGG CTTCAGCGGC CCCGAGTTCA CCTTCCCCCT GTACGGCACC
    IleMetAla SerProVal GlyPheSerGly ProGluPhe ThrPhePro LeuTyrGlyThr
```

Fig. 9B

1021 ATGGGCAACG CTGCACCTCA GCAGCGCATC GTGGCACAGC TGGGCCAGGG AGTGTACCGC
MetGlyAsn AlaAlaPro GlnGlnArgIle ValAlaGln LeuGlyGln GlyValTyrArg

1081 ACCCTGAGCA GCACCCTGTA CCGTCGACCT TTCAACATCG GCATCAACAA CCAGCAGCTG
ThrLeuSer SerThrLeu TyrArgArgPro PheAsnIle GlyIleAsn AsnGlnGlnLeu

1141 AGCGTGCTGG ACGGCACCGA GTTCGCCTAC GGCACCAGCA GCAACCTGCC CAGCGCCGTG
SerValLeu AspGlyThr GluPheAlaTyr GlyThrSer SerAsnLeu ProSerAlaVal

1201 TACCGCAAGA GCGGCACCGT GGACAGCCTG GACGAGATCC CCCCTCAGAA CAACAACGTG
TyrArgLys SerGlyThr ValAspSerLeu AspGluIle ProProGln AsnAsnAsnVal

1261 CCACCTCGAC AGGGCTTCAG CCACCGTCTG AGCCACGTGA GCATGTTCCG CAGTGGCTTC
ProProArg GlnGlyPhe SerHisArgLeu SerHisVal SerMetPhe ArgSerGlyPhe

1321 AGCAACAGCA GCGTGAGCAT CATCCGTGCA CCTATGTTCA GCTGGATTCA CCGCAGTGCC
SerAsnSer SerValSer IleIleArgAla ProMetPhe SerTrpIle HisArgSerAla

1381 GAGTTCAACA ACATCATCCC CAGCAGCCAG ATCACCCAGA TCCCCCTGAC CAAGAGCACC
GluPheAsn AsnIleIle ProSerSerGln IleThrGln IleProLeu ThrLysSerThr

1441 AACCTGGGCA GCGGCACCAG CGTGGTGAAG GGCCCCGGCT TCACCGGCGG CGACATCCTG
AsnLeuGly SerGlyThr SerValValLys GlyProGly PheThrGly GlyAspIleLeu

1501 CGCCGCACCA GCCCCGGCCA GATCAGCACC CTGCGCGTGA ACATCACCGC CCCCCTGAGC
ArgArgThr SerProGly GlnIleSerThr LeuArgVal AsnIleThr AlaProLeuSer

1561 CAGCGCTACC GCGTCCGCAT CCGCTACGCC AGCACCACCA ACCTGCAGTT CCACACCAGC
GlnArgTyr ArgValArg IleArgTyrAla SerThrThr AsnLeuGln PheHisThrSer

1621 ATCGACGGCC GCCCCATCAA CCAGGGCAAC TTCAGCGCCA CCATGAGCAG CGGCAGCAAC
IleAspGly ArgProIle AsnGlnGlyAsn PheSerAla ThrMetSer SerGlySerAsn

1681 CTGCAGAGCG GCAGCTTCCG CACCGTGGGC TTCACCACCC CCTTCAACTT CAGCAACGGC
LeuGlnSer GlySerPhe ArgThrValGly PheThrThr ProPheAsn PheSerAsnGly

1741 AGCAGCGTGT TCACCCTGAG CGCCCACGTG TTCAACAGCG GCAACGAGGT GTACATCGAC
SerSerVal PheThrLeu SerAlaHisVal PheAsnSer GlyAsnGlu ValTyrIleAsp

1801 CGCATCGAGT TCGTGCCCGC CGAGGTGACC TTCGAGGCCG AGTACGACCT GGAGAGGGCT
ArgIleGlu PheValPro AlaGluValThr PheGluAla GluTyrAsp LeuGluArgAla

1861 CAGAAGGCCG TGAACGAGCT GTTCACCAGC AGCAACCAGA TCGGCCTGAA GACCGACGTG
GlnLysAla ValAsnGlu LeuPheThrSer SerAsnGln IleGlyLeu LysThrAspVal

1921 ACCGACTACC ACATCGATCA AGTATCCAAT TTAGTTGAGT GTTTATCTGA TGAATTTTGT
ThrAspTyr HisIleAsp GlnValSerAsn LeuValGlu CysLeuSer AspGluPheCys

1981 CTGGATGAAA AAAAAGAATT GTCCGAGAAA GTCAAACATG CGAAGCGACT TAGTGATGAG
LeuAspGlu LysLysGlu LeuSerGluLys ValLysHis AlaLysArg LeuSerAspGlu

Fig. 9C

```
2041 CGGAATTTAC TTCAAGATCC AAACTTTAGA GGGATCAATA GACAACTAGA CCGTGGCTGG
     ArgAsnLeu LeuGlnAsp ProAsnPheArg GlyIleAsn ArgGlnLeu AspArgGlyTrp

2101 AGAGGAAGTA CGGATATTAC CATCCAAGGA GGCGATGACG TATTCAAAGA GAATTACGTT
     ArgGlySer ThrAspIle ThrIleGlnGly GlyAspAsp ValPheLys GluAsnTyrVal

2161 ACGCTATTGG GTACCTTCGA CGAGTGCTAC CCCACCTACC TGTACCAGAA GATCGACGAG
     ThrLeuLeu GlyThrPhe AspGluCysTyr ProThrTyr LeuTyrGln LysIleAspGlu

2221 AGCAAGCTGA AGGCCTACAC CCGCTACCAG CTGCGCGGCT ACATCGAGGA CAGCCAGGAC
     SerLysLeu LysAlaTyr ThrArgTyrGln LeuArgGly TyrIleGlu AspSerGlnAsp

2281 CTGGAAATCT ACCTGATCCG CTACAACGCC AAGCACGAGA CCGTGAACGT GCCCGGCACC
     LeuGluIle TyrLeuIle ArgTyrAsnAla LysHisGlu ThrValAsn ValProGlyThr

2341 GGCAGCCTGT GGCCCCTGAG CGCCCCCAGC CCCATCGGCA AGTGCGGGGA GCCGAATCGA
     GlySerLeu TrpProLeu SerAlaProSer ProIleGly LysCysGly GluProAsnArg

2401 TGCGCTCCGC ACCTGGAGTG GAACCCGGAC CTAGACTGCA GCTGCAGGGA CGGGGAGAAG
     CysAlaPro HisLeuGlu TrpAsnProAsp LeuAspCys SerCysArg AspGlyGluLys

2461 TGCGCCCACC ACAGCCACCA CTTCAGCCTG GACATCGACG TGGGCTGCAC CGACCTGAAC
     CysAlaHis HisSerHis HisPheSerLeu AspIleAsp ValGlyCys ThrAspLeuAsn

2521 GAGGACCTGG GCGTGTGGGT GATCTTCAAG ATCAAGACCC AGGACGGCCA CGCCCGCCTG
     GluAspLeu GlyValTrp ValIlePheLys IleLysThr GlnAspGly HisAlaArgLeu

2581 GGCAATCTAG AATTTCTCGA AGAGAAACCA TTAGTAGGAG AAGCACTAGC TCGTGTGAAA
     GlyAsnLeu GluPheLeu GluGluLysPro LeuValGly GluAlaLeu AlaArgValLys

2641 AGAGCGGAGA AAAAATGGAG AGACAAACGT GAAAAATTGG AATGGGAAAC AAATATTGTT
     ArgAlaGlu LysLysTrp ArgAspLysArg GluLysLeu GluTrpGlu ThrAsnIleVal

2701 TATAAAGAGG CAAAAGAATC TGTAGATGCT TTATTTGTAA ACTCTCAATA TGATAGATTA
     TyrLysGlu AlaLysGlu SerValAspAla LeuPheVal AsnSerGln TyrAspArgLeu

2761 CAAGCGGATA CCAACATCGC GATGATTCAT GCGGCAGATA AACGCGTTCA TAGCATTCGA
     GlnAlaAsp ThrAsnIle AlaMetIleHis AlaAlaAsp LysArgVal HisSerIleArg

2821 GAAGCTTATC TGCCTGAGCT GTCTGTGATT CCGGGTGTCA ATGCGGCTAT TTTTGAAGAA
     GluAlaTyr LeuProGlu LeuSerValIle ProGlyVal AsnAlaAla IlePheGluGlu

2881 TTAGAAGGGC GTATTTTCAC TGCATTCTCC CTATATGATG CGAGAAATGT CATTAAAAAT
     LeuGluGly ArgIlePhe ThrAlaPheSer LeuTyrAsp AlaArgAsn ValIleLysAsn

2941 GGTGATTTTA ATAATGGCTT ATCCTGCTGG AACGTGAAAG GGCATGTAGA TGTAGAAGAA
     GlyAspPhe AsnAsnGly LeuSerCysTrp AsnValLys GlyHisVal AspValGluGlu

3001 CAAAACAACC ACCGTTCGGT CCTTGTTGTT CCGGAATGGG AAGCAGAAGT GTCACAAGAA
     GlnAsnAsn HisArgSer ValLeuValVal ProGluTrp GluAlaGlu ValSerGlnGlu
```

Fig. 9D

```
3061 GTTCGTGTCT GTCCGGGTCG TGGCTATATC CTTCGTGTCA CAGCGTACAA GGAGGGATAT
     ValArgVal CysProGly ArgGlyTyrIle LeuArgVal ThrAlaTyr LysGluGlyTyr

3121 GGAGAAGGTT GCGTAACCAT TCATGAGATC GAGAACAATA CAGACGAACT GAAGTTTAGC
     GlyGluGly CysValThr IleHisGluIle GluAsnAsn ThrAspGlu LeuLysPheSer

3181 AACTGTGTAG AAGAGGAAGT ATATCCAAAC AACACGGTAA CGTGTAATGA TTATACTGCG
     AsnCysVal GluGluGlu ValTyrProAsn AsnThrVal ThrCysAsn AspTyrThrAla

3241 ACTCAAGAAG AATATGAGGG TACGTACACT TCTCGTAATC GAGGATATGA CGGAGCCTAT
     ThrGlnGlu GluTyrGlu GlyThrTyrThr SerArgAsn ArgGlyTyr AspGlyAlaTyr

3301 GAAAGCAATT CTTCTGTACC AGCTGATTAT GCATCAGCCT ATGAAGAAAA AGCATATACA
     GluSerAsn SerSerVal ProAlaAspTyr AlaSerAla TyrGluGlu LysAlaTyrThr

3361 GATGGACGAA GAGACAATCC TTGTGAATCT AACAGAGGAT ATGGGGATTA CACACCACTA
     AspGlyArg ArgAspAsn ProCysGluSer AsnArgGly TyrGlyAsp TyrThrProLeu

3421 CCAGCTGGCT ATGTGACAAA AGAATTAGAG TACTTCCCAG AAACCGATAA GGTATGGATT
     ProAlaGly TyrValThr LysGluLeuGlu TyrPhePro GluThrAsp LysValTrpIle

3481 GAGATCGGAG AAACGGAAGG AACATTCATC GTGGACAGCG TGGAATTACT TCTTATGGAG
     GluIleGly GluThrGlu GlyThrPheIle ValAspSer ValGluLeu LeuLeuMetGlu

3541 GAATAA
     Glu---
```

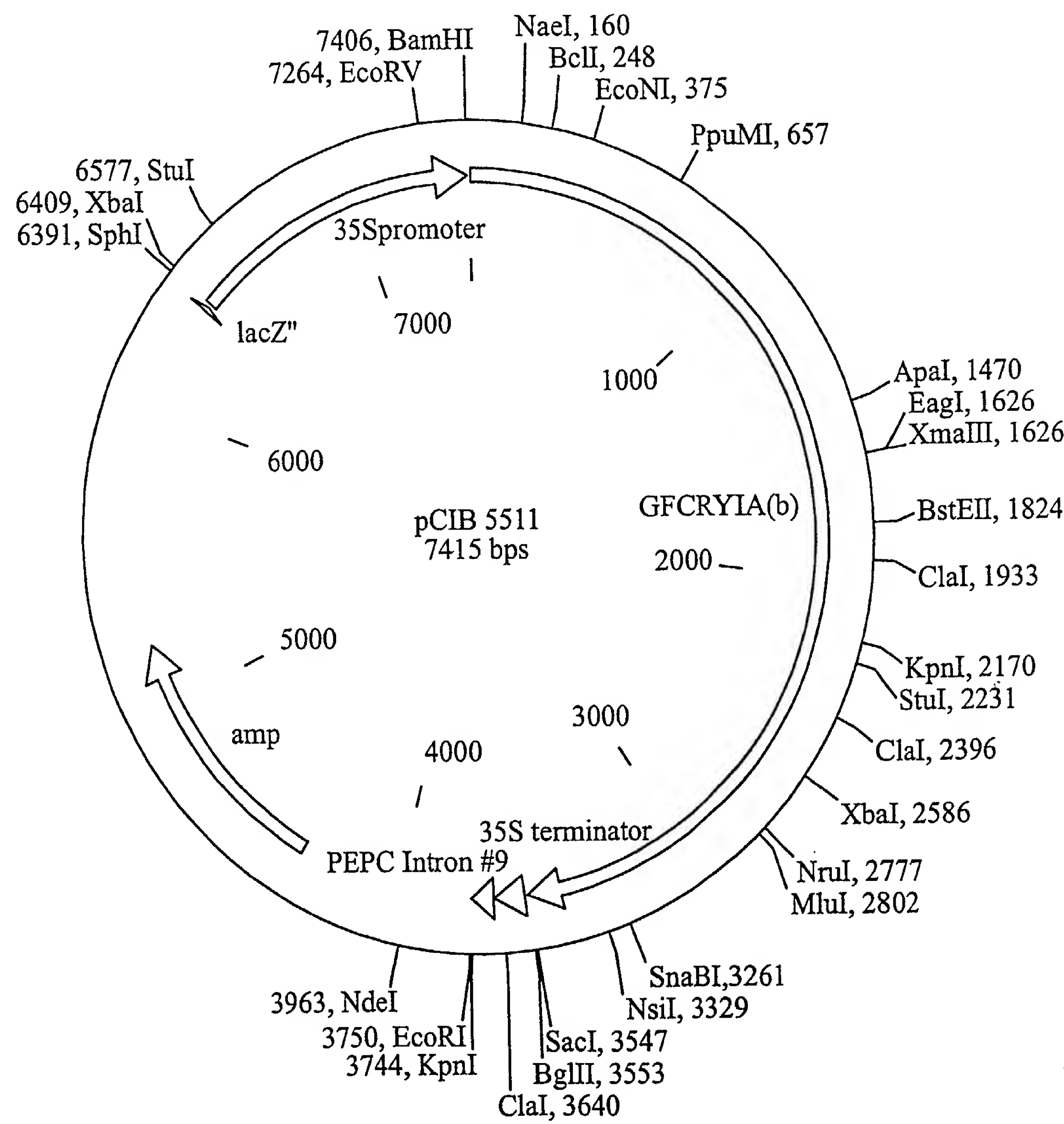

7406, BamHI
7264, EcoRV
NaeI, 160
BclI, 248
EcoNI, 375
PpuMI, 657
6577, StuI
6409, XbaI
6391, SphI
35Spromoter
lacZ"
7000
1000
ApaI, 1470
EagI, 1626
XmaIII, 1626
6000
pCIB 5511
7415 bps
GFCRYIA(b)
BstEII, 1824
2000
ClaI, 1933
5000
KpnI, 2170
StuI, 2231
amp
3000
4000
ClaI, 2396
XbaI, 2586
35S terminator
PEPC Intron #9
NruI, 2777
MluI, 2802
SnaBI,3261
NsiI, 3329
3963, NdeI
SacI, 3547
3750, EcoRI
3744, KpnI
BglII, 3553
ClaI, 3640

Fig. 10

Fig. 11A

1 ATGGACAACA ACCCCAACAT CAACGAGTGC ATCCCCTACA ACTGCCTGAG CAACCCCGAG
MetAspAsn AsnProAsn IleAsnGluCys IleProTyr AsnCysLeu SerAsnProGlu

61 GTGGAGGTGC TGGGCGGCGA GCGCATCGAG ACCGGCTACA CCCCCATCGA CATCAGCCTG
ValGluVal LeuGlyGly GluArgIleGlu ThrGlyTyr ThrProIle AspIleSerLeu

121 AGCCTGACCC AGTTCCTGCT GAGCGAGTTC GTGCCCGGCG CCGGCTTCGT GCTGGGCCTG
SerLeuThr GlnPheLeu LeuSerGluPhe ValProGly AlaGlyPhe ValLeuGlyLeu

181 GTGGACATCA TCTGGGGCAT CTTCGGCCCC AGCCAGTGGG ACGCCTTCCT GGTGCAGATC
ValAspIle IleTrpGly IlePheGlyPro SerGlnTrp AspAlaPhe LeuValGlnIle

241 GAGCAGCTGA TCAACCAGCG CATCGAGGAG TTCGCCCGCA ACCAGGCCAT CAGCCGCCTG
GluGlnLeu IleAsnGln ArgIleGluGlu PheAlaArg AsnGlnAla IleSerArgLeu

301 GAGGGCCTGA GCAACCTGTA CCAAATCTAC GCCGAGAGCT TCCGCGAGTG GGAGGCCGAC
GluGlyLeu SerAsnLeu TyrGlnIleTyr AlaGluSer PheArgGlu TrpGluAlaAsp

361 CCCACCAACC CCGCCCTGCG CGAGGAGATG CGCATCCAGT TCAACGACAT GAACAGCGCC
ProThrAsn ProAlaLeu ArgGluGluMet ArgIleGln PheAsnAsp MetAsnSerAla

421 CTGACCACCG CCATCCCCCT GTTCGCCGTG CAGAACTACC AGGTGCCCCT GCTGAGCGTG
LeuThrThr AlaIlePro LeuPheAlaVal GlnAsnTyr GlnValPro LeuLeuSerVal

481 TACGTGCAGG CCGCCAACCT GCACCTGAGC GTGCTGCGCG ACGTCAGCGT GTTCGGCCAG
TyrValGln AlaAlaAsn LeuHisLeuSer ValLeuArg AspValSer ValPheGlyGln

541 CGCTGGGGCT TCGACGCCGC CACCATCAAC AGCCGCTACA ACGACCTGAC CCGCCTGATC
ArgTrpGly PheAspAla AlaThrIleAsn SerArgTyr AsnAspLeu ThrArgLeuIle

601 GGCAACTACA CCGACCACGC CGTGCGCTGG TACAACACCG GCCTGGAGCG CGTGTGGGGT
GlyAsnTyr ThrAspHis AlaValArgTrp TyrAsnThr GlyLeuGlu ArgValTrpGly

661 CCCGACAGCC GCGACTGGAT CAGGTACAAC CAGTTCCGCC GCGAGCTGAC CCTGACCGTG
ProAspSer ArgAspTrp IleArgTyrAsn GlnPheArg ArgGluLeu ThrLeuThrVal

721 CTGGACATCG TGAGCCTGTT CCCCAACTAC GACAGCCGCA CCTACCCCAT CCGCACCGTG
LeuAspIle ValSerLeu PheProAsnTyr AspSerArg ThrTyrPro IleArgThrVal

781 AGCCAGCTGA CCCGCGAGAT TTACACCAAC CCCGTGCTGG AGAACTTCGA CGGCAGCTTC
SerGlnLeu ThrArgGlu IleTyrThrAsn ProValLeu GluAsnPhe AspGlySerPhe

841 CGCGGCAGCG CCCAGGGCAT CGAGGGCAGC ATCCGCAGCC CCCACCTGAT GGACATCCTG
ArgGlySer AlaGlnGly IleGluGlySer IleArgSer ProHisLeu MetAspIleLeu

901 AACAGCATCA CCATCTACAC CGACGCCCAC CGCGGCGAGT ACTACTGGAG CGGCCACCAG
AsnSerIle ThrIleTyr ThrAspAlaHis ArgGlyGlu TyrTyrTrp SerGlyHisGln

961 ATCATGGCCA GCCCCGTCGG CTTCAGCGGC CCCGAGTTCA CCTTCCCCCT GTACGGCACC
IleMetAla SerProVal GlyPheSerGly ProGluPhe ThrPhePro LeuTyrGlyThr

Fig. 11B

```
1021  ATGGGCAACG CTGCACCTCA GCAGCGCATC GTGGCACAGC TGGGCCAGGG AGTGTACCGC
      MetGlyAsn AlaAlaPro GlnGlnArgIle ValAlaGln LeuGlyGln GlyValTyrArg

1081  ACCCTGAGCA GCACCCTGTA CCGTCGACCT TTCAACATCG GCATCAACAA CCAGCAGCTG
      ThrLeuSer SerThrLeu TyrArgArgPro PheAsnIle GlyIleAsn AsnGlnGlnLeu

1141  AGCGTGCTGG ACGGCACCGA GTTCGCCTAC GGCACCAGCA GCAACCTGCC CAGCGCCGTG
      SerValLeu AspGlyThr GluPheAlaTyr GlyThrSer SerAsnLeu ProSerAlaVal

1201  TACCGCAAGA GCGGCACCGT GGACAGCCTG GACGAGATCC CCCCTCAGAA CAACAACGTG
      TyrArgLys SerGlyThr ValAspSerLeu AspGluIle ProProGln AsnAsnAsnVal

1261  CCACCTCGAC AGGGCTTCAG CCACCGTCTG AGCCACGTGA GCATGTTCCG CAGTGGCTTC
      ProProArg GlnGlyPhe SerHisArgLeu SerHisVal SerMetPhe ArgSerGlyPhe

1321  AGCAACAGCA GCGTGAGCAT CATCCGTGCA CCTATGTTCA GCTGGATTCA CCGCAGTGCC
      SerAsnSer SerValSer IleIleArgAla ProMetPhe SerTrpIle HisArgSerAla

1381  GAGTTCAACA ACATCATCCC CAGCAGCCAG ATCACCCAGA TCCCCCTGAC CAAGAGCACC
      GluPheAsn AsnIleIle ProSerSerGln IleThrGln IleProLeu ThrLysSerThr

1441  AACCTGGGCA GCGGCACCAG CGTGGTGAAG GGCCCCGGCT TCACCGGCGG CGACATCCTG
      AsnLeuGly SerGlyThr SerValValLys GlyProGly PheThrGly GlyAspIleLeu

1501  CGCCGCACCA GCCCCGGCCA GATCAGCACC CTGCGCGTGA ACATCACCGC CCCCCTGAGC
      ArgArgThr SerProGly GlnIleSerThr LeuArgVal AsnIleThr AlaProLeuSer

1561  CAGCGCTACC GCGTCCGCAT CCGCTACGCC AGCACCACCA ACCTGCAGTT CCACACCAGC
      GlnArgTyr ArgValArg IleArgTyrAla SerThrThr AsnLeuGln PheHisThrSer

1621  ATCGACGGCC GCCCCATCAA CCAGGGCAAC TTCAGCGCCA CCATGAGCAG CGGCAGCAAC
      IleAspGly ArgProIle AsnGlnGlyAsn PheSerAla ThrMetSer SerGlySerAsn

1681  CTGCAGAGCG GCAGCTTCCG CACCGTGGGC TTCACCACCC CCTTCAACTT CAGCAACGGC
      LeuGlnSer GlySerPhe ArgThrValGly PheThrThr ProPheAsn PheSerAsnGly

1741  AGCAGCGTGT TCACCCTGAG CGCCCACGTG TTCAACAGCG GCAACGAGGT GTACATCGAC
      SerSerVal PheThrLeu SerAlaHisVal PheAsnSer GlyAsnGlu ValTyrIleAsp

1801  CGCATCGAGT TCGTGCCCGC CGAGGTGACC TTCGAGGCCG AGTACGACCT GGAGAGGGCT
      ArgIleGlu PheValPro AlaGluValThr PheGluAla GluTyrAsp LeuGluArgAla

1861  CAGAAGGCCG TGAACGAGCT GTTCACCAGC AGCAACCAGA TCGGCCTGAA GACCGACGTG
      GlnLysAla ValAsnGlu LeuPheThrSer SerAsnGln IleGlyLeu LysThrAspVal

1921  ACCGACTACC ACATCGATCA GGTGAGCAAC CTGGTGGAGT GCTTAAGCGA CGAGTTCTGC
      ThrAspTyr HisIleAsp GlnValSerAsn LeuValGlu CysLeuSer AspGluPheCys

1981  CTGGACGAGA AGAAGGAGCT GAGCGAGAAG GTGAAGCACG CCAAGCGCCT GAGCGACGAG
      LeuAspGlu LysLysGlu LeuSerGluLys ValLysHis AlaLysArg LeuSerAspGlu
```

Fig. 11C

```
2041  CGCAACCTGC TGCAGGACCC CAACTTCCGC GGCATCAACC GCCAGCTGGA CCGCGGCTGG
      ArgAsnLeu LeuGlnAsp ProAsnPheArg GlyIleAsn ArgGlnLeu AspArgGlyTrp

2101  CGAGGCAGCA CCGATATCAC CATCCAGGGC GGCGACGACG TGTTCAAGGA GAACTACGTG
      ArgGlySer ThrAspIle ThrIleGlnGly GlyAspAsp ValPheLys GluAsnTyrVal

2161  ACCCTGCTGG GCACCTTCGA CGAGTGCTAC CCCACCTACC TGTACCAGAA GATCGACGAG
      ThrLeuLeu GlyThrPhe AspGluCysTyr ProThrTyr LeuTyrGln LysIleAspGlu

2221  AGCAAGCTGA AGGCCTACAC CCGCTACCAG CTGCGCGGCT ACATCGAGGA CAGCCAGGAC
      SerLysLeu LysAlaTyr ThrArgTyrGln LeuArgGly TyrIleGlu AspSerGlnAsp

2281  CTGGAAATCT ACCTGATCCG CTACAACGCC AAGCACGAGA CCGTGAACGT GCCCGGCACC
      LeuGluIle TyrLeuIle ArgTyrAsnAla LysHisGlu ThrValAsn ValProGlyThr

2341  GGCAGCCTGT GGCCCCTGAG CGCCCCCAGC CCCATCGGCA AGTGCGGGGA GCCGAATCGA
      GlySerLeu TrpProLeu SerAlaProSer ProIleGly LysCysGly GluProAsnArg

2401  TGCGCTCCGC ACCTGGAGTG GAACCCGGAC CTAGACTGCA GCTGCAGGGA CGGGGAGAAG
      CysAlaPro HisLeuGlu TrpAsnProAsp LeuAspCys SerCysArg AspGlyGluLys

2461  TGCGCCCACC ACAGCCACCA CTTCAGCCTG GACATCGACG TGGGCTGCAC CGACCTGAAC
      CysAlaHis HisSerHis HisPheSerLeu AspIleAsp ValGlyCys ThrAspLeuAsn

2521  GAGGACCTGG GCGTGTGGGT GATCTTCAAG ATCAAGACCC AGGACGGCCA CGCCCGCCTG
      GluAspLeu GlyValTrp ValIlePheLys IleLysThr GlnAspGly HisAlaArgLeu

2581  GGCAATCTAG AATTTCTCGA AGAGAAACCA TTAGTAGGAG AAGCACTAGC TCGTGTGAAA
      GlyAsnLeu GluPheLeu GluGluLysPro LeuValGly GluAlaLeu AlaArgValLys

2641  AGAGCGGAGA AAAAATGGAG AGACAAACGT GAAAAATTGG AATGGGAAAC AAATATTGTT
      ArgAlaGlu LysLysTrp ArgAspLysArg GluLysLeu GluTrpGlu ThrAsnIleVal

2701  TATAAAGAGG CAAAAGAATC TGTAGATGCT TTATTTGTAA ACTCTCAATA TGATAGATTA
      TyrLysGlu AlaLysGlu SerValAspAla LeuPheVal AsnSerGln TyrAspArgLeu

2761  CAAGCGGATA CCAACATCGC GATGATTCAT GCGGCAGATA AACGCGTTCA TAGCATTCGA
      GlnAlaAsp ThrAsnIle AlaMetIleHis AlaAlaAsp LysArgVal HisSerIleArg

2821  GAAGCTTATC TGCCTGAGCT GTCTGTGATT CCGGGTGTCA ATGCGGCTAT TTTTGAAGAA
      GluAlaTyr LeuProGlu LeuSerValIle ProGlyVal AsnAlaAla IlePheGluGlu

2881  TTAGAAGGGC GTATTTTCAC TGCATTCTCC CTATATGATG CGAGAAATGT CATTAAAAAT
      LeuGluGly ArgIlePhe ThrAlaPheSer LeuTyrAsp AlaArgAsn ValIleLysAsn

2941  GGTGATTTTA ATAATGGCTT ATCCTGCTGG AACGTGAAAG GCATGTAGA TGTAGAAGAA
      GlyAspPhe AsnAsnGly LeuSerCysTrp AsnValLys GlyHisVal AspValGluGlu

3001  CAAAACAACC ACCGTTCGGT CCTTGTTGTT CCGGAATGGG AAGCAGAAGT GTCACAAGAA
      GlnAsnAsn HisArgSer ValLeuValVal ProGluTrp GluAlaGlu ValSerGlnGlu
```

Fig. 11D

```
3061  GTTCGTGTCT GTCCGGGTCG TGGCTATATC CTTCGTGTCA CAGCGTACAA GGAGGGATAT
      ValArgVal CysProGly ArgGlyTyrIle LeuArgVal ThrAlaTyr LysGluGlyTyr

3121  GGAGAAGGTT GCGTAACCAT TCATGAGATC GAGAACAATA CAGACGAACT GAAGTTTAGC
      GlyGluGly CysValThr IleHisGluIle GluAsnAsn ThrAspGlu LeuLysPheSer

3181  AACTGTGTAG AAGAGGAAGT ATATCCAAAC AACACGGTAA CGTGTAATGA TTATACTGCG
      AsnCysVal GluGluGlu ValTyrProAsn AsnThrVal ThrCysAsn AspTyrThrAla

3241  ACTCAAGAAG AATATGAGGG TACGTACACT TCTCGTAATC GAGGATATGA CGGAGCCTAT
      ThrGlnGlu GluTyrGlu GlyThrTyrThr SerArgAsn ArgGlyTyr AspGlyAlaTyr

3301  GAAAGCAATT CTTCTGTACC AGCTGATTAT GCATCAGCCT ATGAAGAAAA AGCATATACA
      GluSerAsn SerSerVal ProAlaAspTyr AlaSerAla TyrGluGlu LysAlaTyrThr

3361  GATGGACGAA GAGACAATCC TTGTGAATCT AACAGAGGAT ATGGGGATTA CACACCACTA
      AspGlyArg ArgAspAsn ProCysGluSer AsnArgGly TyrGlyAsp TyrThrProLeu

3421  CCAGCTGGCT ATGTGACAAA AGAATTAGAG TACTTCCCAG AAACCGATAA GGTATGGATT
      ProAlaGly TyrValThr LysGluLeuGlu TyrPhePro GluThrAsp LysValTrpIle

3481  GAGATCGGAG AAACGGAAGG AACATTCATC GTGGACAGCG TGGAATTACT TCTTATGGAG
      GluIleGly GluThrGlu GlyThrPheIle ValAspSer ValGluLeu LeuLeuMetGlu

3541  GAATAA
      Glu---
```

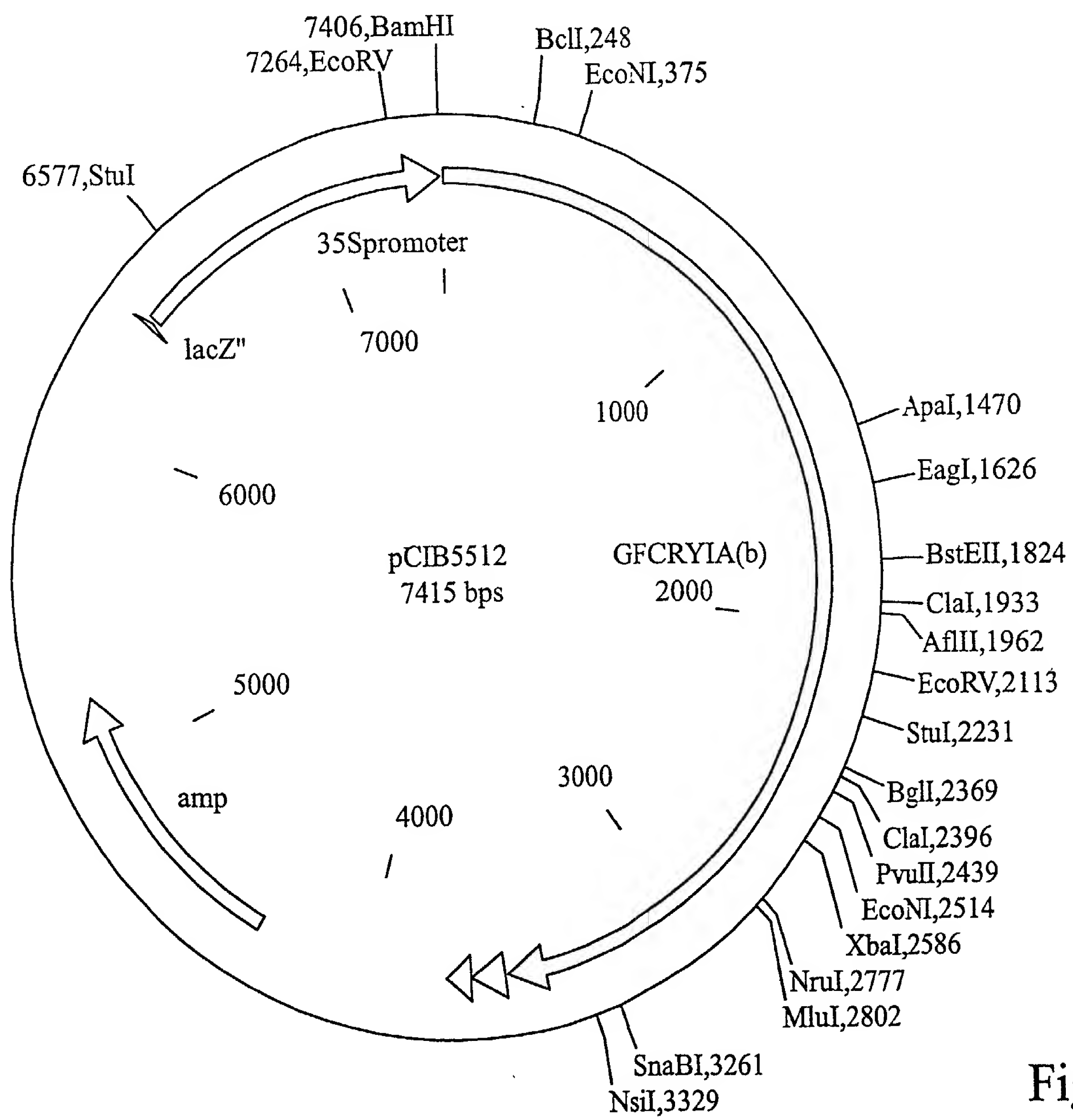

7406,BamHI
7264,EcoRV
BclI,248
EcoNI,375
6577,StuI
35Spromoter
lacZ"
7000
1000
6000
pCIB5512
7415 bps
GFCRYIA(b)
2000
5000
amp
4000
3000
ApaI,1470
EagI,1626
BstEII,1824
ClaI,1933
AflII,1962
EcoRV,2113
StuI,2231
BglI,2369
ClaI,2396
PvuII,2439
EcoNI,2514
XbaI,2586
NruI,2777
MluI,2802
SnaBI,3261
NsiI,3329

Fig. 12

Fig. 13A

1 ATGGACAACA ACCCCAACAT CAACGAGTGC ATCCCCTACA ACTGCCTGAG CAACCCCGAG
MetAspAsn AsnProAsn IleAsnGluCys IleProTyr AsnCysLeu SerAsnProGlu

61 GTGGAGGTGC TGGGCGGCGA GCGCATCGAG ACCGGCTACA CCCCCATCGA CATCAGCCTG
ValGluVal LeuGlyGly GluArgIleGlu ThrGlyTyr ThrProIle AspIleSerLeu

121 AGCCTGACCC AGTTCCTGCT GAGCGAGTTC GTGCCCGGCG CCGGCTTCGT GCTGGGCCTG
SerLeuThr GlnPheLeu LeuSerGluPhe ValProGly AlaGlyPhe ValLeuGlyLeu

181 GTGGACATCA TCTGGGGCAT CTTCGGCCCC AGCCAGTGGG ACGCCTTCCT GGTGCAGATC
ValAspIle IleTrpGly IlePheGlyPro SerGlnTrp AspAlaPhe LeuValGlnIle

241 GAGCAGCTGA TCAACCAGCG CATCGAGGAG TTCGCCCGCA ACCAGGCCAT CAGCCGCCTG
GluGlnLeu IleAsnGln ArgIleGluGlu PheAlaArg AsnGlnAla IleSerArgLeu

301 GAGGGCCTGA GCAACCTGTA CCAAATCTAC GCCGAGAGCT TCCGCGAGTG GGAGGCCGAC
GluGlyLeu SerAsnLeu TyrGlnIleTyr AlaGluSer PheArgGlu TrpGluAlaAsp

361 CCCACCAACC CCGCCCTGCG CGAGGAGATG CGCATCCAGT TCAACGACAT GAACAGCGCC
ProThrAsn ProAlaLeu ArgGluGluMet ArgIleGln PheAsnAsp MetAsnSerAla

421 CTGACCACCG CCATCCCCCT GTTCGCCGTG CAGAACTACC AGGTGCCCCT GCTGAGCGTG
LeuThrThr AlaIlePro LeuPheAlaVal GlnAsnTyr GlnValPro LeuLeuSerVal

481 TACGTGCAGG CCGCCAACCT GCACCTGAGC GTGCTGCGCG ACGTCAGCGT GTTCGGCCAG
TyrValGln AlaAlaAsn LeuHisLeuSer ValLeuArg AspValSer ValPheGlyGln

541 CGCTGGGGCT TCGACGCCGC CACCATCAAC AGCCGCTACA ACGACCTGAC CCGCCTGATC
ArgTrpGly PheAspAla AlaThrIleAsn SerArgTyr AsnAspLeu ThrArgLeuIle

601 GGCAACTACA CCGACCACGC CGTGCGCTGG TACAACACCG GCCTGGAGCG CGTGTGGGGT
GlyAsnTyr ThrAspHis AlaValArgTrp TyrAsnThr GlyLeuGlu ArgValTrpGly

661 CCCGACAGCC GCGACTGGAT CAGGTACAAC CAGTTCCGCC GCGAGCTGAC CCTGACCGTG
ProAspSer ArgAspTrp IleArgTyrAsn GlnPheArg ArgGluLeu ThrLeuThrVal

721 CTGGACATCG TGAGCCTGTT CCCCAACTAC GACAGCCGCA CCTACCCCAT CCGCACCGTG
LeuAspIle ValSerLeu PheProAsnTyr AspSerArg ThrTyrPro IleArgThrVal

781 AGCCAGCTGA CCCGCGAGAT TTACACCAAC CCCGTGCTGG AGAACTTCGA CGGCAGCTTC
SerGlnLeu ThrArgGlu IleTyrThrAsn ProValLeu GluAsnPhe AspGlySerPhe

841 CGCGGCAGCG CCCAGGGCAT CGAGGGCAGC ATCCGCAGCC CCCACCTGAT GGACATCCTG
ArgGlySer AlaGlnGly IleGluGlySer IleArgSer ProHisLeu MetAspIleLeu

901 AACAGCATCA CCATCTACAC CGACGCCCAC CGCGGCGAGT ACTACTGGAG CGGCCACCAG
AsnSerIle ThrIleTyr ThrAspAlaHis ArgGlyGlu TyrTyrTrp SerGlyHisGln

961 ATCATGGCCA GCCCCGTCGG CTTCAGCGGC CCCGAGTTCA CCTTCCCCCT GTACGGCACC
IleMetAla SerProVal GlyPheSerGly ProGluPhe ThrPhePro LeuTyrGlyThr

Fig. 13B

1021 ATGGGCAACG CTGCACCTCA GCAGCGCATC GTGGCACAGC TGGGCCAGGG AGTGTACCGC
MetGlyAsn AlaAlaPro GlnGlnArgIle ValAlaGln LeuGlyGln GlyValTyrArg

1081 ACCCTGAGCA GCACCCTGTA CCGTCGACCT TTCAACATCG GCATCAACAA CCAGCAGCTG
ThrLeuSer SerThrLeu TyrArgArgPro PheAsnIle GlyIleAsn AsnGlnGlnLeu

1141 AGCGTGCTGG ACGGCACCGA GTTCGCCTAC GGCACCAGCA GCAACCTGCC CAGCGCCGTG
SerValLeu AspGlyThr GluPheAlaTyr GlyThrSer SerAsnLeu ProSerAlaVal

1201 TACCGCAAGA GCGGCACCGT GGACAGCCTG GACGAGATCC CCCCTCAGAA CAACAACGTG
TyrArgLys SerGlyThr ValAspSerLeu AspGluIle ProProGln AsnAsnAsnVal

1261 CCACCTCGAC AGGGCTTCAG CCACCGTCTG AGCCACGTGA GCATGTTCCG CAGTGGCTTC
ProProArg GlnGlyPhe SerHisArgLeu SerHisVal SerMetPhe ArgSerGlyPhe

1321 AGCAACAGCA GCGTGAGCAT CATCCGTGCA CCTATGTTCA GCTGGATTCA CCGCAGTGCC
SerAsnSer SerValSer IleIleArgAla ProMetPhe SerTrpIle HisArgSerAla

1381 GAGTTCAACA ACATCATCCC CAGCAGCCAG ATCACCCAGA TCCCCCTGAC CAAGAGCACC
GluPheAsn AsnIleIle ProSerSerGln IleThrGln IleProLeu ThrLysSerThr

1441 AACCTGGGCA GCGGCACCAG CGTGGTGAAG GGCCCCGGCT TCACCGGCGG CGACATCCTG
AsnLeuGly SerGlyThr SerValValLys GlyProGly PheThrGly GlyAspIleLeu

1501 CGCCGCACCA GCCCCGGCCA GATCAGCACC CTGCGCGTGA ACATCACCGC CCCCCTGAGC
ArgArgThr SerProGly GlnIleSerThr LeuArgVal AsnIleThr AlaProLeuSer

1561 CAGCGCTACC GCGTCCGCAT CCGCTACGCC AGCACCACCA ACCTGCAGTT CCACACCAGC
GlnArgTyr ArgValArg IleArgTyrAla SerThrThr AsnLeuGln PheHisThrSer

1621 ATCGACGGCC GCCCCATCAA CCAGGGCAAC TTCAGCGCCA CCATGAGCAG CGGCAGCAAC
IleAspGly ArgProIle AsnGlnGlyAsn PheSerAla ThrMetSer SerGlySerAsn

1681 CTGCAGAGCG GCAGCTTCCG CACCGTGGGC TTCACCACCC CCTTCAACTT CAGCAACGGC
LeuGlnSer GlySerPhe ArgThrValGly PheThrThr ProPheAsn PheSerAsnGly

1741 AGCAGCGTGT TCACCCTGAG CGCCCACGTG TTCAACAGCG GCAACGAGGT GTACATCGAC
SerSerVal PheThrLeu SerAlaHisVal PheAsnSer GlyAsnGlu ValTyrIleAsp

1801 CGCATCGAGT TCGTGCCCGC CGAGGTGACC TTCGAGGCCG AGTACGACCT GGAGAGGGCT
ArgIleGlu PheValPro AlaGluValThr PheGluAla GluTyrAsp LeuGluArgAla

1861 CAGAAGGCCG TGAACGAGCT GTTCACCAGC AGCAACCAGA TCGGCCTGAA GACCGACGTG
GlnLysAla ValAsnGlu LeuPheThrSer SerAsnGln IleGlyLeu LysThrAspVal

1921 ACCGACTACC ACATCGATCA GGTGAGCAAC CTGGTGGAGT GCTTAAGCGA CGAGTTCTGC
ThrAspTyr HisIleAsp GlnValSerAsn LeuValGlu CysLeuSer AspGluPheCys

1981 CTGGACGAGA AGAAGGAGCT GAGCGAGAAG GTGAAGCACG CCAAGCGCCT GAGCGACGAG
LeuAspGlu LysLysGlu LeuSerGluLys ValLysHis AlaLysArg LeuSerAspGlu

Fig. 13C

```
2041  CGCAACCTGC TGCAGGACCC CAACTTCCGC GGCATCAACC GCCAGCTGGA CCGCGGCTGG
      ArgAsnLeu LeuGlnAsp ProAsnPheArg GlyIleAsn ArgGlnLeu AspArgGlyTrp

2101  CGAGGCAGCA CCGATATCAC CATCCAGGGC GGCGACGACG TGTTCAAGGA GAACTACGTG
      ArgGlySer ThrAspIle ThrIleGlnGly GlyAspAsp ValPheLys GluAsnTyrVal

2161  ACCCTGCTGG GCACCTTCGA CGAGTGCTAC CCCACCTACC TGTACCAGAA GATCGACGAG
      ThrLeuLeu GlyThrPhe AspGluCysTyr ProThrTyr LeuTyrGln LysIleAspGlu

2221  AGCAAGCTGA AGGCCTACAC CCGCTACCAG CTGCGCGGCT ACATCGAGGA CAGCCAGGAC
      SerLysLeu LysAlaTyr ThrArgTyrGln LeuArgGly TyrIleGlu AspSerGlnAsp

2281  CTGGAAATCT ACCTGATCCG CTACAACGCC AAGCACGAGA CCGTGAACGT GCCCGGCACC
      LeuGluIle TyrLeuIle ArgTyrAsnAla LysHisGlu ThrValAsn ValProGlyThr

2341  GGCAGCCTGT GGCCCCTGAG CGCCCCCAGC CCCATCGGCA AGTGCGGGGA GCCGAATCGA
      GlySerLeu TrpProLeu SerAlaProSer ProIleGly LysCysGly GluProAsnArg

2401  TGCGCTCCGC ACCTGGAGTG GAACCCGGAC CTAGACTGCA GCTGCAGGGA CGGGGAGAAG
      CysAlaPro HisLeuGlu TrpAsnProAsp LeuAspCys SerCysArg AspGlyGluLys

2461  TGCGCCCACC ACAGCCACCA CTTCAGCCTG GACATCGACG TGGGCTGCAC CGACCTGAAC
      CysAlaHis HisSerHis HisPheSerLeu AspIleAsp ValGlyCys ThrAspLeuAsn

2521  GAGGACCTGG GCGTGTGGGT GATCTTCAAG ATCAAGACCC AGGACGGCCA CGCCCGCCTG
      GluAspLeu GlyValTrp ValIlePheLys IleLysThr GlnAspGly HisAlaArgLeu

2581  GGCAATCTAG AGTTCCTGGA GGAGAAGCCC CTGGTGGGCG AGGCCCTGGC CCGCGTGAAG
      GlyAsnLeu GluPheLeu GluGluLysPro LeuValGly GluAlaLeu AlaArgValLys

2641  CGCGCCGAGA AGAAGTGGCG CGACAAGCGC GAGAAGCTGG AGTGGGAGAC CAACATCGTG
      ArgAlaGlu LysLysTrp ArgAspLysArg GluLysLeu GluTrpGlu ThrAsnIleVal

2701  TACAAGGAGG CCAAGGAGAG CGTGGACGCC CTGTTCGTGA ACAGCCAGTA CGACCGCCTG
      TyrLysGlu AlaLysGlu SerValAspAla LeuPheVal AsnSerGln TyrAspArgLeu

2761  CAGGCCGACA CCAACATCGC CATGATCCAC GCCGCCGACA AGCGCGTGCA CAGCATTCGC
      GlnAlaAsp ThrAsnIle AlaMetIleHis AlaAlaAsp LysArgVal HisSerIleArg

2821  GAGGCCTACC TGCCCGAGCT GAGCGTGATC CCCGGCGTGA ACGCCGCCAT CTTCGAGGAA
      GluAlaTyr LeuProGlu LeuSerValIle ProGlyVal AsnAlaAla IlePheGluGlu

2881  CTCGAGGGCC GCATCTTCAC CGCCTTCAGC CTGTACGACG CCCGCAACGT GATCAAGAAC
      LeuGluGly ArgIlePhe ThrAlaPheSer LeuTyrAsp AlaArgAsn ValIleLysAsn

2941  GGCGACTTCA ACAACGGCCT GAGCTGCTGG AACGTGAAGG GCCACGTGGA CGTGGAGGAG
      GlyAspPhe AsnAsnGly LeuSerCysTrp AsnValLys GlyHisVal AspValGluGlu

3001  CAGAACAACC ACCGCAGCGT GCTGGTGGTG CCCGAGTGGG AGGCCGAGGT GAGCCAGGAG
      GlnAsnAsn HisArgSer ValLeuValVal ProGluTrp GluAlaGlu ValSerGlnGlu
```

Fig. 13D

```
3061  GTGCGCGTGT GCCCCGGCCG CGGCTACATC CTGCGCGTGA CCGCCTACAA GGAGGGCTAC
      ValArgVal CysProGly ArgGlyTyrIle LeuArgVal ThrAlaTyr LysGluGlyTyr

3121  GGCGAGGGCT GCGTGACCAT CCACGAGATC GAGAACAACA CCGACGAGCT CAAGTTCAGC
      GlyGluGly CysValThr IleHisGluIle GluAsnAsn ThrAspGlu LeuLysPheSer

3181  AACTGCGTGG AGGAGGAGGT GTACCCCAAC AACACCGTGA CCTGCAACGA CTACACCGCG
      AsnCysVal GluGluGlu ValTyrProAsn AsnThrVal ThrCysAsn AspTyrThrAla

3241  ACCCAGGAGG AGTACGAGGG CACCTACACC AGCCGCAACC GCGGCTACGA CGGCGCCTAC
      ThrGlnGlu GluTyrGlu GlyThrTyrThr SerArgAsn ArgGlyTyr AspGlyAlaTyr

3301  GAGAGCAACA GCAGCGTGCC CGCCGACTAC GCCAGCGCCT ACGAGGAGAA GGCCTACACC
      GluSerAsn SerSerVal ProAlaAspTyr AlaSerAla TyrGluGlu LysAlaTyrThr

3361  GACGGCCGCC GCGACAACCC CTGCGAGAGC AACCGCGGCT ACGGCGACTA CACCCCCCTG
      AspGlyArg ArgAspAsn ProCysGluSer AsnArgGly TyrGlyAsp TyrThrProLeu

3421  CCCGCCGGCT ACGTGACCAA GGAGCTGGAG TACTTCCCCG AGACCGACAA GGTGTGGATC
      ProAlaGly TyrValThr LysGluLeuGlu TyrPhePro GluThrAsp LysValTrpIle

3481  GAGATCGGCG AGACCGAGGG CACCTTCATC GTGGACAGCG TGGAGCTGCT GCTGATGGAG
      GluIleGly GluThrGlu GlyThrPheIle ValAspSer ValGluLeu LeuLeuMetGlu

3541  GAG
      Glu
```

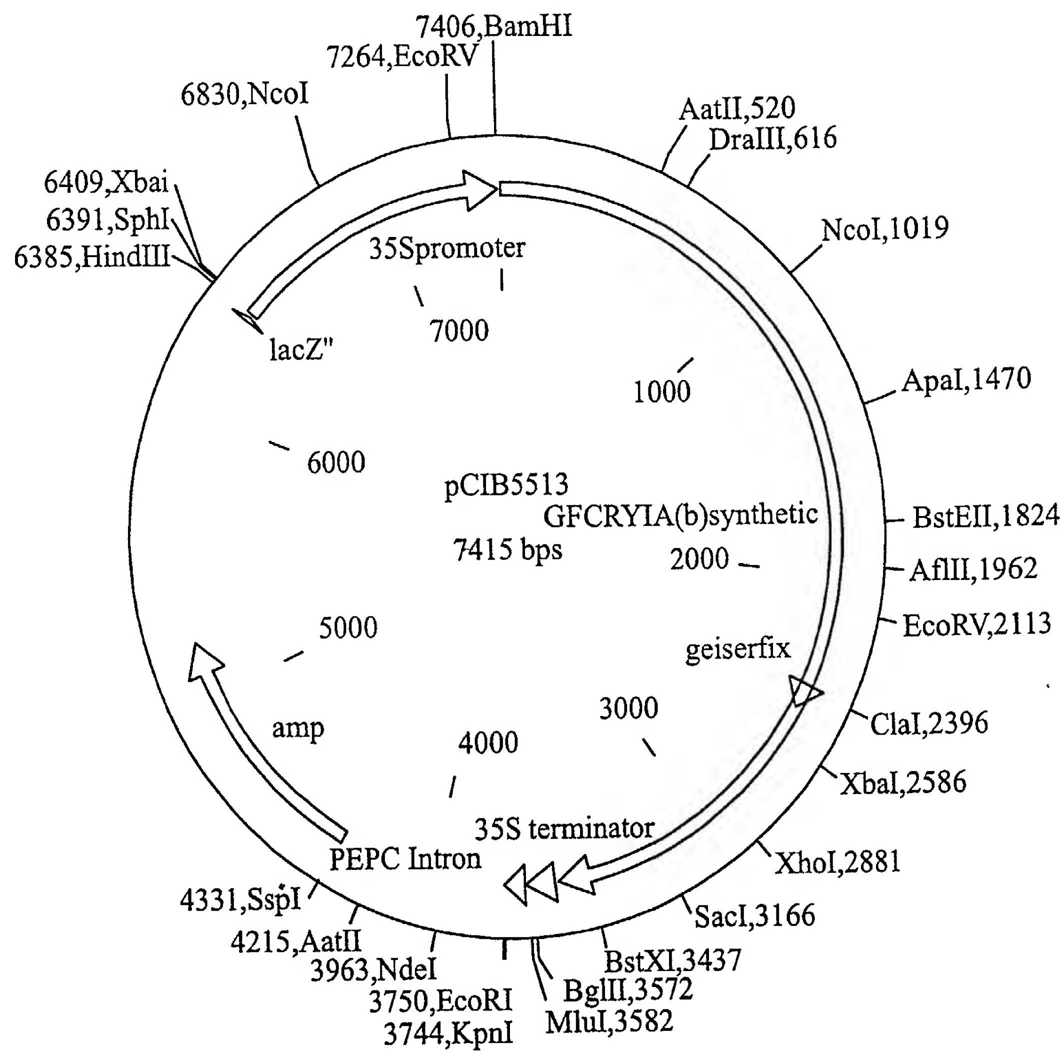

7406,BamHI
7264,EcoRV
6830,NcoI
6409,Xbai
6391,SphI
6385,HindIII
AatII,520
DraIII,616
NcoI,1019
ApaI,1470
BstEII,1824
AflII,1962
EcoRV,2113
ClaI,2396
XbaI,2586
XhoI,2881
SacI,3166
BstXI,3437
BglII,3572
MluI,3582
3744,KpnI
3750,EcoRI
3963,NdeI
4215,AatII
4331,SspI
35Spromoter
lacZ"
7000
1000
6000
pCIB5513
GFCRYIA(b)synthetic
7415 bps
2000
5000
geiserfix
3000
4000
amp
35S terminator
PEPC Intron

Fig. 14

Fig. 15A

```
  1  ATGGACAACA ACCCCAACAT CAACGAGTGC ATCCCCTACA ACTGCCTGAG CAACCCCGAG
     MetAspAsn AsnProAsn IleAsnGluCys IleProTyr AsnCysLeu SerAsnProGlu

 61  GTGGAGGTGC TGGGCGGCGA GCGCATCGAG ACCGGCTACA CCCCCATCGA CATCAGCCTG
     ValGluVal LeuGlyGly GluArgIleGlu ThrGlyTyr ThrProIle AspIleSerLeu

121  AGCCTGACCC AGTTCCTGCT GAGCGAGTTC GTGCCCGGCG CCGGCTTCGT GCTGGGCCTG
     SerLeuThr GlnPheLeu LeuSerGluPhe ValProGly AlaGlyPhe ValLeuGlyLeu

181  GTGGACATCA TCTGGGGCAT CTTCGGCCCC AGCCAGTGGG ACGCCTTCCT GGTGCAGATC
     ValAspIle IleTrpGly IlePheGlyPro SerGlnTrp AspAlaPhe LeuValGlnIle

241  GAGCAGCTGA TCAACCAGCG CATCGAGGAG TTCGCCCGCA ACCAGGCCAT CAGCCGCCTG
     GluGlnLeu IleAsnGln ArgIleGluGlu PheAlaArg AsnGlnAla IleSerArgLeu

301  GAGGGCCTGA GCAACCTGTA CCAAATCTAC GCCGAGAGCT TCCGCGAGTG GGAGGCCGAC
     GluGlyLeu SerAsnLeu TyrGlnIleTyr AlaGluSer PheArgGlu TrpGluAlaAsp

361  CCCACCAACC CCGCCCTGCG CGAGGAGATG CGCATCCAGT TCAACGACAT GAACAGCGCC
     ProThrAsn ProAlaLeu ArgGluGluMet ArgIleGln PheAsnAsp MetAsnSerAla

421  CTGACCACCG CCATCCCCCT GTTCGCCGTG CAGAACTACC AGGTGCCCCT GCTGAGCGTG
     LeuThrThr AlaIlePro LeuPheAlaVal GlnAsnTyr GlnValPro LeuLeuSerVal

481  TACGTGCAGG CCGCCAACCT GCACCTGAGC GTGCTGCGCG ACGTCAGCGT GTTCGGCCAG
     TyrValGln AlaAlaAsn LeuHisLeuSer ValLeuArg AspValSer ValPheGlyGln

541  CGCTGGGGCT TCGACGCCGC CACCATCAAC AGCCGCTACA ACGACCTGAC CCGCCTGATC
     ArgTrpGly PheAspAla AlaThrIleAsn SerArgTyr AsnAspLeu ThrArgLeuIle

601  GGCAACTACA CCGACCACGC CGTGCGCTGG TACAACACCG GCCTGGAGCG CGTGTGGGGT
     GlyAsnTyr ThrAspHis AlaValArgTrp TyrAsnThr GlyLeuGlu ArgValTrpGly

661  CCCGACAGCC GCGACTGGAT CAGGTACAAC CAGTTCCGCC GCGAGCTGAC CCTGACCGTG
     ProAspSer ArgAspTrp IleArgTyrAsn GlnPheArg ArgGluLeu ThrLeuThrVal

721  CTGGACATCG TGAGCCTGTT CCCCAACTAC GACAGCCGCA CCTACCCCAT CCGCACCGTG
     LeuAspIle ValSerLeu PheProAsnTyr AspSerArg ThrTyrPro IleArgThrVal

781  AGCCAGCTGA CCCGCGAGAT TTACACCAAC CCCGTGCTGG AGAACTTCGA CGGCAGCTTC
     SerGlnLeu ThrArgGlu IleTyrThrAsn ProValLeu GluAsnPhe AspGlySerPhe

841  CGCGGCAGCG CCCAGGGCAT CGAGGGCAGC ATCCGCAGCC CCCACCTGAT GGACATCCTG
     ArgGlySer AlaGlnGly IleGluGlySer IleArgSer ProHisLeu MetAspIleLeu

901  AACAGCATCA CCATCTACAC CGACGCCCAC CGCGGCGAGT ACTACTGGAG CGGCCACCAG
     AsnSerIle ThrIleTyr ThrAspAlaHis ArgGlyGlu TyrTyrTrp SerGlyHisGln

961  ATCATGGCCA GCCCCGTCGG CTTCAGCGGC CCCGAGTTCA CCTTCCCCCT GTACGGCACC
     IleMetAla SerProVal GlyPheSerGly ProGluPhe ThrPhePro LeuTyrGlyThr
```

Fig. 15B

```
1021  ATGGGCAACG CTGCACCTCA GCAGCGCATC GTGGCACAGC TGGGCCAGGG AGTGTACCGC
      MetGlyAsn AlaAlaPro GlnGlnArgIle ValAlaGln LeuGlyGln GlyValTyrArg

1081  ACCCTGAGCA GCACCCTGTA CCGTCGACCT TTCAACATCG GCATCAACAA CCAGCAGCTG
      ThrLeuSer SerThrLeu TyrArgArgPro PheAsnIle GlyIleAsn AsnGlnGlnLeu

1141  AGCGTGCTGG ACGGCACCGA GTTCGCCTAC GGCACCAGCA GCAACCTGCC CAGCGCCGTG
      SerValLeu AspGlyThr GluPheAlaTyr GlyThrSer SerAsnLeu ProSerAlaVal

1201  TACCGCAAGA GCGGCACCGT GGACAGCCTG GACGAGATCC CCCCTCAGAA CAACAACGTG
      TyrArgLys SerGlyThr ValAspSerLeu AspGluIle ProProGln AsnAsnAsnVal

1261  CCACCTCGAC AGGGCTTCAG CCACCGTCTG AGCCACGTGA GCATGTTCCG CAGTGGCTTC
      ProProArg GlnGlyPhe SerHisArgLeu SerHisVal SerMetPhe ArgSerGlyPhe

1321  AGCAACAGCA GCGTGAGCAT CATCCGTGCA CCTATGTTCA GCTGGATTCA CCGCAGTGCC
      SerAsnSer SerValSer IleIleArgAla ProMetPhe SerTrpIle HisArgSerAla

1381  GAGTTCAACA ACATCATCCC CAGCAGCCAG ATCACCCAGA TCCCCCTGAC CAAGAGCACC
      GluPheAsn AsnIleIle ProSerSerGln IleThrGln IleProLeu ThrLysSerThr

1441  AACCTGGGCA GCGGCACCAG CGTGGTGAAG GGCCCCGGCT TCACCGGCGG CGACATCCTG
      AsnLeuGly SerGlyThr SerValValLys GlyProGly PheThrGly GlyAspIleLeu

1501  CGCCGCACCA GCCCCGGCCA GATCAGCACC CTGCGCGTGA ACATCACCGC CCCCCTGAGC
      ArgArgThr SerProGly GlnIleSerThr LeuArgVal AsnIleThr AlaProLeuSer

1561  CAGCGCTACC GCGTCCGCAT CCGCTACGCC AGCACCACCA ACCTGCAGTT CCACACCAGC
      GlnArgTyr ArgValArg IleArgTyrAla SerThrThr AsnLeuGln PheHisThrSer

1621  ATCGACGGCC GCCCCATCAA CCAGGGCAAC TTCAGCGCCA CCATGAGCAG CGGCAGCAAC
      IleAspGly ArgProIle AsnGlnGlyAsn PheSerAla ThrMetSer SerGlySerAsn

1681  CTGCAGAGCG GCAGCTTCCG CACCGTGGGC TTCACCACCC CCTTCAACTT CAGCAACGGC
      LeuGlnSer GlySerPhe ArgThrValGly PheThrThr ProPheAsn PheSerAsnGly

1741  AGCAGCGTGT TCACCCTGAG CGCCCACGTG TTCAACAGCG GCAACGAGGT GTACATCGAC
      SerSerVal PheThrLeu SerAlaHisVal PheAsnSer GlyAsnGlu ValTyrIleAsp

1801  CGCATCGAGT TCGTGCCCGC CGAGGTGACC TTCGAGGCCG AGTACGACCT GGAGAGGGCT
      ArgIleGlu PheValPro AlaGluValThr PheGluAla GluTyrAsp LeuGluArgAla

1861  CAGAAGGCCG TGAACGAGCT GTTCACCAGC AGCAACCAGA TCGGCCTGAA GACCGACGTG
      GlnLysAla ValAsnGlu LeuPheThrSer SerAsnGln IleGlyLeu LysThrAspVal

1921  ACCGACTACC ACATCGATCA AGTATCCAAT TTAGTTGAGT GTTTATCTGA TGAATTTTGT
      ThrAspTyr HisIleAsp GlnValSerAsn LeuValGlu CysLeuSer AspGluPheCys

1981  CTGGATGAAA AAAAAGAATT GTCCGAGAAA GTCAAACATG CGAAGCGACT TAGTGATGAG
      LeuAspGlu LysLysGlu LeuSerGluLys ValLysHis AlaLysArg LeuSerAspGlu
```

Fig. 15C

2041 CGGAATTTAC TTCAAGATCC AAACTTTAGA GGGATCAATA GACAACTAGA CCGTGGCTGG
ArgAsnLeu LeuGlnAsp ProAsnPheArg GlyIleAsn ArgGlnLeu AspArgGlyTrp

2101 AGAGGAAGTA CGGATATTAC CATCCAAGGA GGCGATGACG TATTCAAAGA GAATTACGTT
ArgGlySer ThrAspIle ThrIleGlnGly GlyAspAsp ValPheLys GluAsnTyrVal

2161 ACGCTATTGG GTACCTTTGA TGAGTGCTAT CCAACGTATT TATATCAAAA AATAGATGAG
ThrLeuLeu GlyThrPhe AspGluCysTyr ProThrTyr LeuTyrGln LysIleAspGlu

2221 TCGAAATTAA AAGCCTATAC CCGTTACCAA TTAAGAGGGT ATATCGAAGA TAGTCAAGAC
SerLysLeu LysAlaTyr ThrArgTyrGln LeuArgGly TyrIleGlu AspSerGlnAsp

2281 TTAGAAATCT ATTTAATTCG CTACAATGCC AAACACGAAA CAGTAAATGT GCCAGGTACG
LeuGluIle TyrLeuIle ArgTyrAsnAla LysHisGlu ThrValAsn ValProGlyThr

2341 GGTTCCTTAT GGCCGCTTTC AGCCCCAAGT CCAATCGGCA AGTGCGGGGA GCCGAATCGA
GlySerLeu TrpProLeu SerAlaProSer ProIleGly LysCysGly GluProAsnArg

2401 TGCGCTCCGC ACCTGGAGTG GAACCCGGAC CTAGACTGCA GCTGCAGGGA CGGGGAGAAG
CysAlaPro HisLeuGlu TrpAsnProAsp LeuAspCys SerCysArg AspGlyGluLys

2461 TGCGCCCACC ACAGCCACCA CTTCAGCCTG GACATCGACG TGGGCTGCAC CGACCTGAAC
CysAlaHis HisSerHis HisPheSerLeu AspIleAsp ValGlyCys ThrAspLeuAsn

2521 GAGGACCTGG GCGTGTGGGT GATCTTCAAG ATCAAGACCC AGGACGGCCA CGCCCGCCTG
GluAspLeu GlyValTrp ValIlePheLys IleLysThr GlnAspGly HisAlaArgLeu

2581 GGCAATCTAG AATTTCTCGA AGAGAAACCA TTAGTAGGAG AAGCACTAGC TCGTGTGAAA
GlyAsnLeu GluPheLeu GluGluLysPro LeuValGly GluAlaLeu AlaArgValLys

2641 AGAGCGGAGA AAAAATGGAG AGACAAACGT GAAAAATTGG AATGGGAAAC AAATATTGTT
ArgAlaGlu LysLysTrp ArgAspLysArg GluLysLeu GluTrpGlu ThrAsnIleVal

2701 TATAAAGAGG CAAAAGAATC TGTAGATGCT TTATTTGTAA ACTCTCAATA TGATAGATTA
TyrLysGlu AlaLysGlu SerValAspAla LeuPheVal AsnSerGln TyrAspArgLeu

2761 CAAGCGGATA CCAACATCGC GATGATTCAT GCGGCAGATA AACGCGTTCA TAGCATTCGA
GlnAlaAsp ThrAsnIle AlaMetIleHis AlaAlaAsp LysArgVal HisSerIleArg

2821 GAAGCTTATC TGCCTGAGCT GTCTGTGATT CCGGGTGTCA ATGCGGCTAT TTTTGAAGAA
GluAlaTyr LeuProGlu LeuSerValIle ProGlyVal AsnAlaAla IlePheGluGlu

2881 TTAGAAGGGC GTATTTTCAC TGCATTCTCC CTATATGATG CGAGAAATGT CATTAAAAAT
LeuGluGly ArgIlePhe ThrAlaPheSer LeuTyrAsp AlaArgAsn ValIleLysAsn

2941 GGTGATTTTA ATAATGGCTT ATCCTGCTGG AACGTGAAAG GGCATGTAGA TGTAGAAGAA
GlyAspPhe AsnAsnGly LeuSerCysTrp AsnValLys GlyHisVal AspValGluGlu

3001 CAAAACAACC ACCGTTCGGT CCTTGTTGTT CCGGAATGGG AAGCAGAAGT GTCACAAGAA
GlnAsnAsn HisArgSer ValLeuValVal ProGluTrp GluAlaGlu ValSerGlnGlu

Fig. 15D

```
3061  GTTCGTGTCT GTCCGGGTCG TGGCTATATC CTTCGTGTCA CAGCGTACAA GGAGGGATAT
      ValArgVal CysProGly ArgGlyTyrIle LeuArgVal ThrAlaTyr LysGluGlyTyr

3121  GGAGAAGGTT GCGTAACCAT TCATGAGATC GAGAACAATA CAGACGAACT GAAGTTTAGC
      GlyGluGly CysValThr IleHisGluIle GluAsnAsn ThrAspGlu LeuLysPheSer

3181  AACTGTGTAG AAGAGGAAGT ATATCCAAAC AACACGGTAA CGTGTAATGA TTATACTGCG
      AsnCysVal GluGluGlu ValTyrProAsn AsnThrVal ThrCysAsn AspTyrThrAla

3241  ACTCAAGAAG AATATGAGGG TACGTACACT TCTCGTAATC GAGGATATGA CGGAGCCTAT
      ThrGlnGlu GluTyrGlu GlyThrTyrThr SerArgAsn ArgGlyTyr AspGlyAlaTyr

3301  GAAAGCAATT CTTCTGTACC AGCTGATTAT GCATCAGCCT ATGAAGAAAA AGCATATACA
      GluSerAsn SerSerVal ProAlaAspTyr AlaSerAla TyrGluGlu LysAlaTyrThr

3361  GATGGACGAA GAGACAATCC TTGTGAATCT AACAGAGGAT ATGGGGATTA CACACCACTA
      AspGlyArg ArgAspAsn ProCysGluSer AsnArgGly TyrGlyAsp TyrThrProLeu

3421  CCAGCTGGCT ATGTGACAAA AGAATTAGAG TACTTCCCAG AAACCGATAA GGTATGGATT
      ProAlaGly TyrValThr LysGluLeuGlu TyrPhePro GluThrAsp LysValTrpIle

3481  GAGATCGGAG AAACGGAAGG AACATTCATC GTGGACAGCG TGGAATTACT TCTTATGGAG
      GluIleGly GluThrGlu GlyThrPheIle ValAspSer ValGluLeu LeuLeuMetGlu

3541  GAATAAG
      Glu---
```

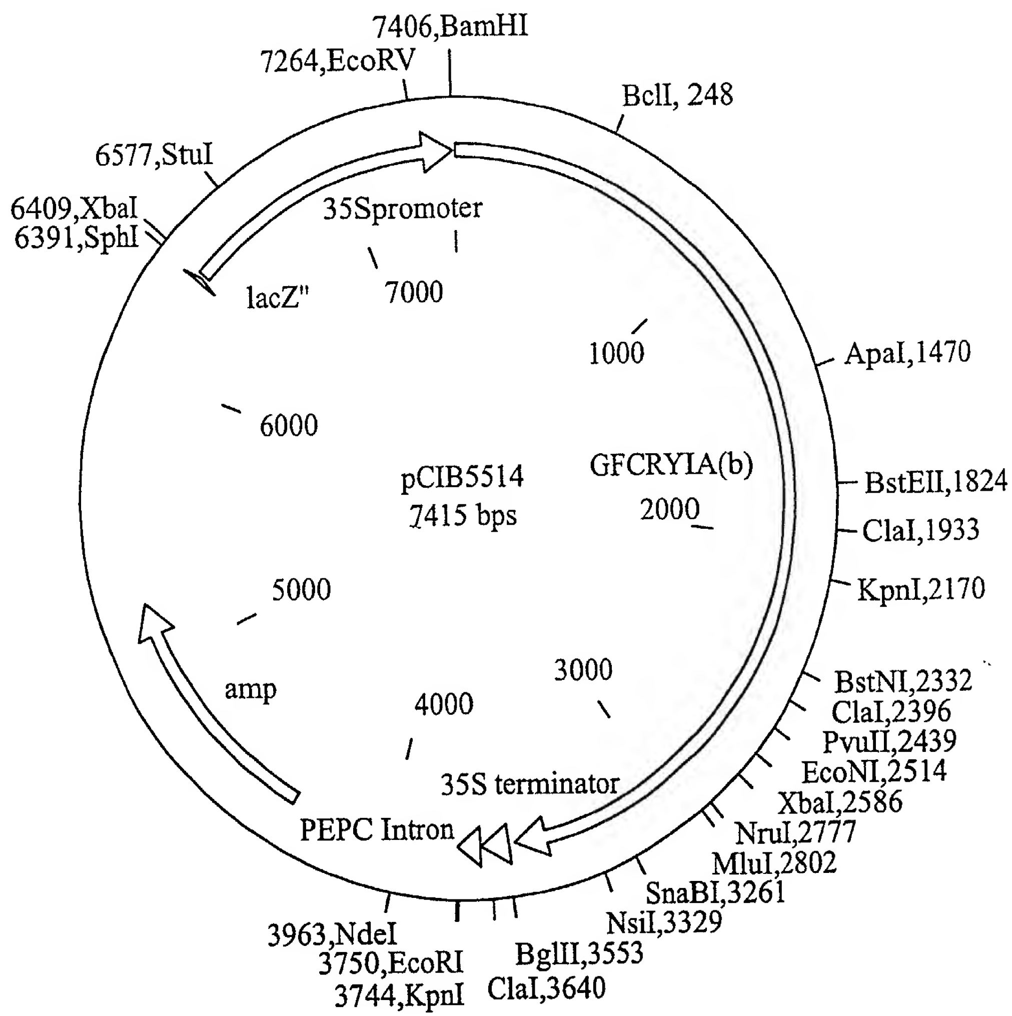

7406,BamHI
7264,EcoRV
BclI, 248
6577,StuI
6409,XbaI
6391,SphI
35Spromoter
lacZ"
7000
1000
ApaI,1470
6000
pCIB5514
7415 bps
GFCRYIA(b)
2000
BstEII,1824
ClaI,1933
KpnI,2170
5000
3000
4000
amp
BstNI,2332
ClaI,2396
PvuII,2439
EcoNI,2514
XbaI,2586
NruI,2777
MluI,2802
SnaBI,3261
NsiI,3329
BglII,3553
ClaI,3640
35S terminator
PEPC Intron
3963,NdeI
3750,EcoRI
3744,KpnI

Fig. 16

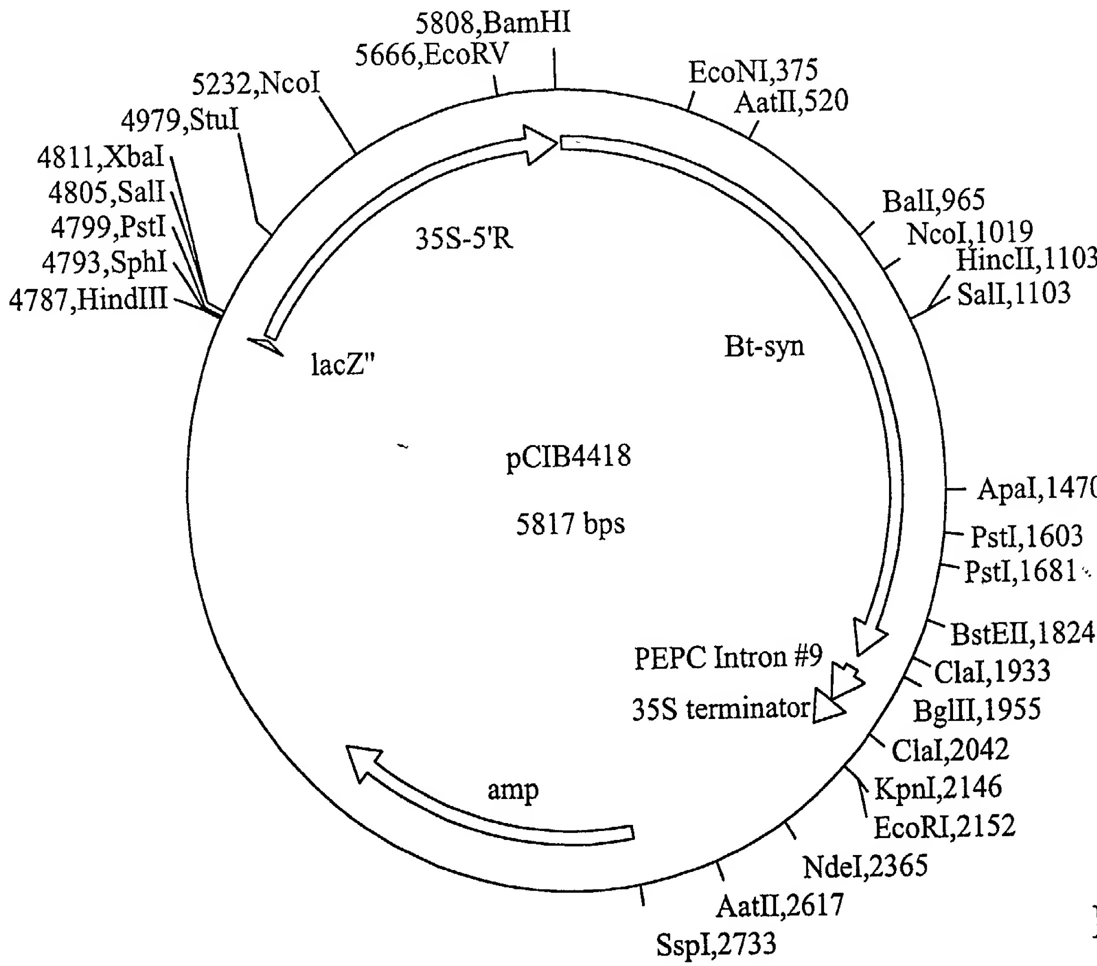

5808,BamHI
5666,EcoRV
5232,NcoI
4979,StuI
4811,XbaI
4805,SalI
4799,PstI
4793,SphI
4787,HindIII
EcoNI,375
AatII,520
BalI,965
NcoI,1019
HincII,1103
SalI,1103
ApaI,1470
PstI,1603
PstI,1681
BstEII,1824
ClaI,1933
BglII,1955
ClaI,2042
KpnI,2146
EcoRI,2152
NdeI,2365
AatII,2617
SspI,2733
35S-5'R
lacZ"
Bt-syn
pCIB4418
5817 bps
PEPC Intron #9
35S terminator
amp

Fig. 17

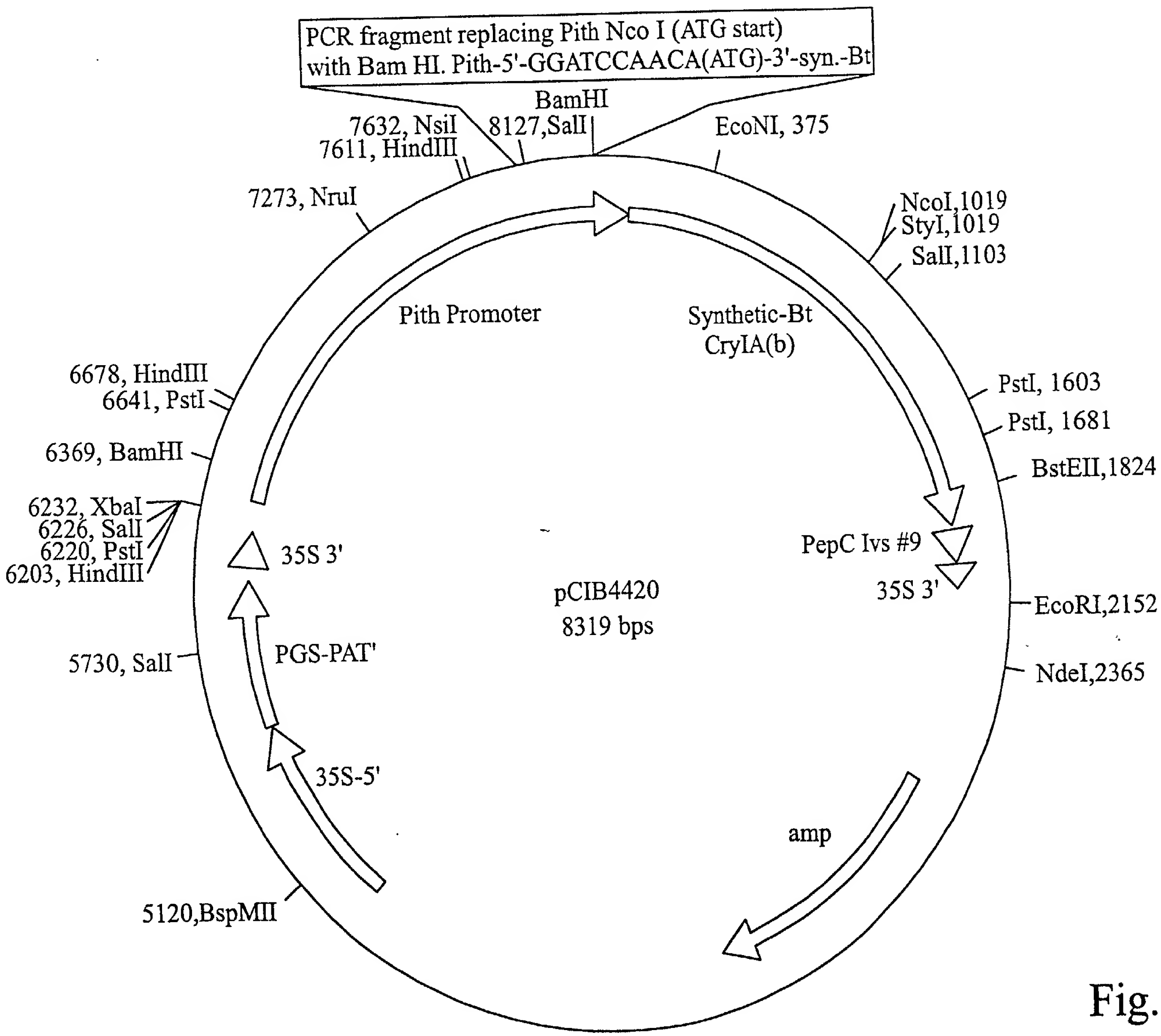

PCR fragment replacing Pith Nco I (ATG start)
with Bam HI. Pith-5'-GGATCCAACA(ATG)-3'-syn.-Bt
BamHI
8127,SalI
7632, NsiI
7611, HindIII
EcoNI, 375
7273, NruI
NcoI,1019
StyI,1019
SalI,1103
Pith Promoter
Synthetic-Bt
CryIA(b)
PstI, 1603
PstI, 1681
BstEII,1824
6678, HindIII
6641, PstI
6369, BamHI
6232, XbaI
6226, SalI
6220, PstI
6203, HindIII
35S 3'
PepC Ivs #9
35S 3'
pCIB4420
8319 bps
EcoRI,2152
5730, SalI
PGS-PAT'
NdeI,2365
35S-5'
amp
5120,BspMII

Fig. 18

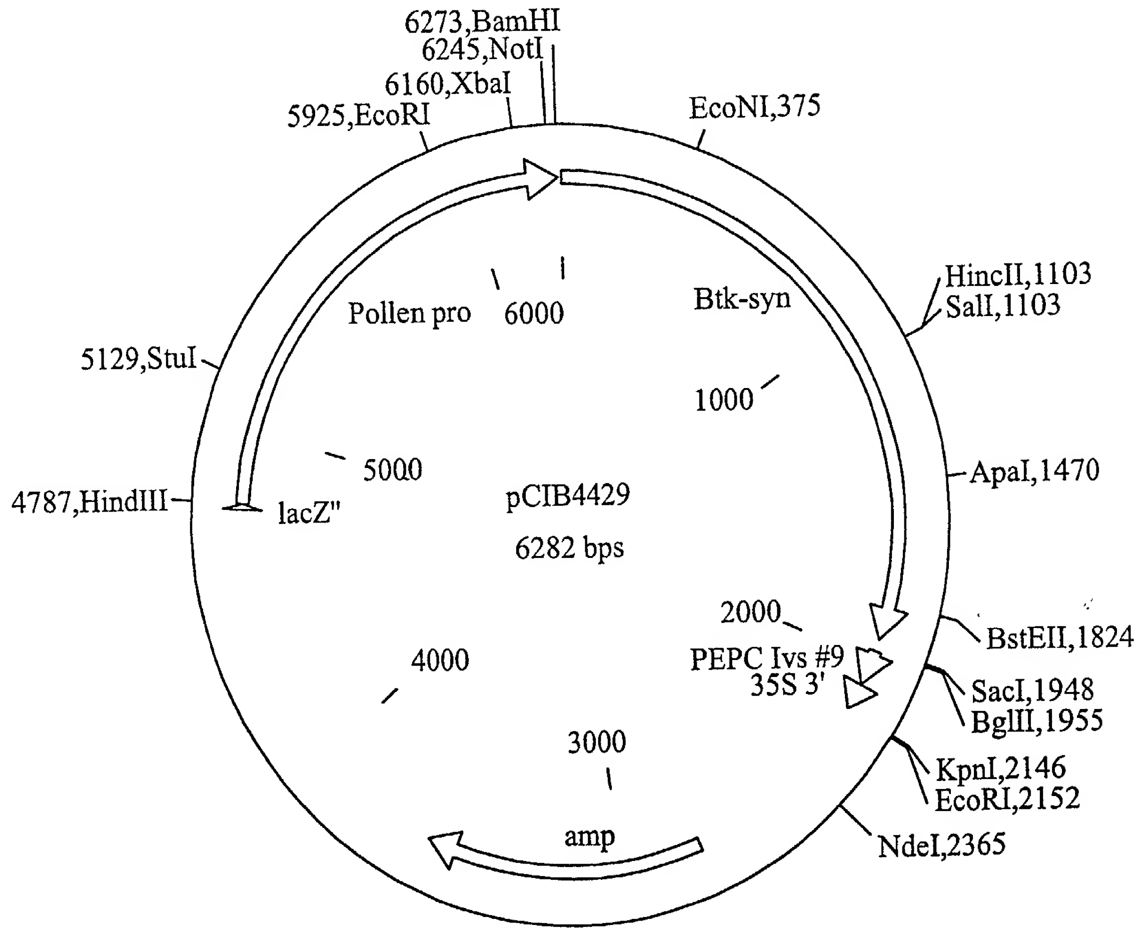
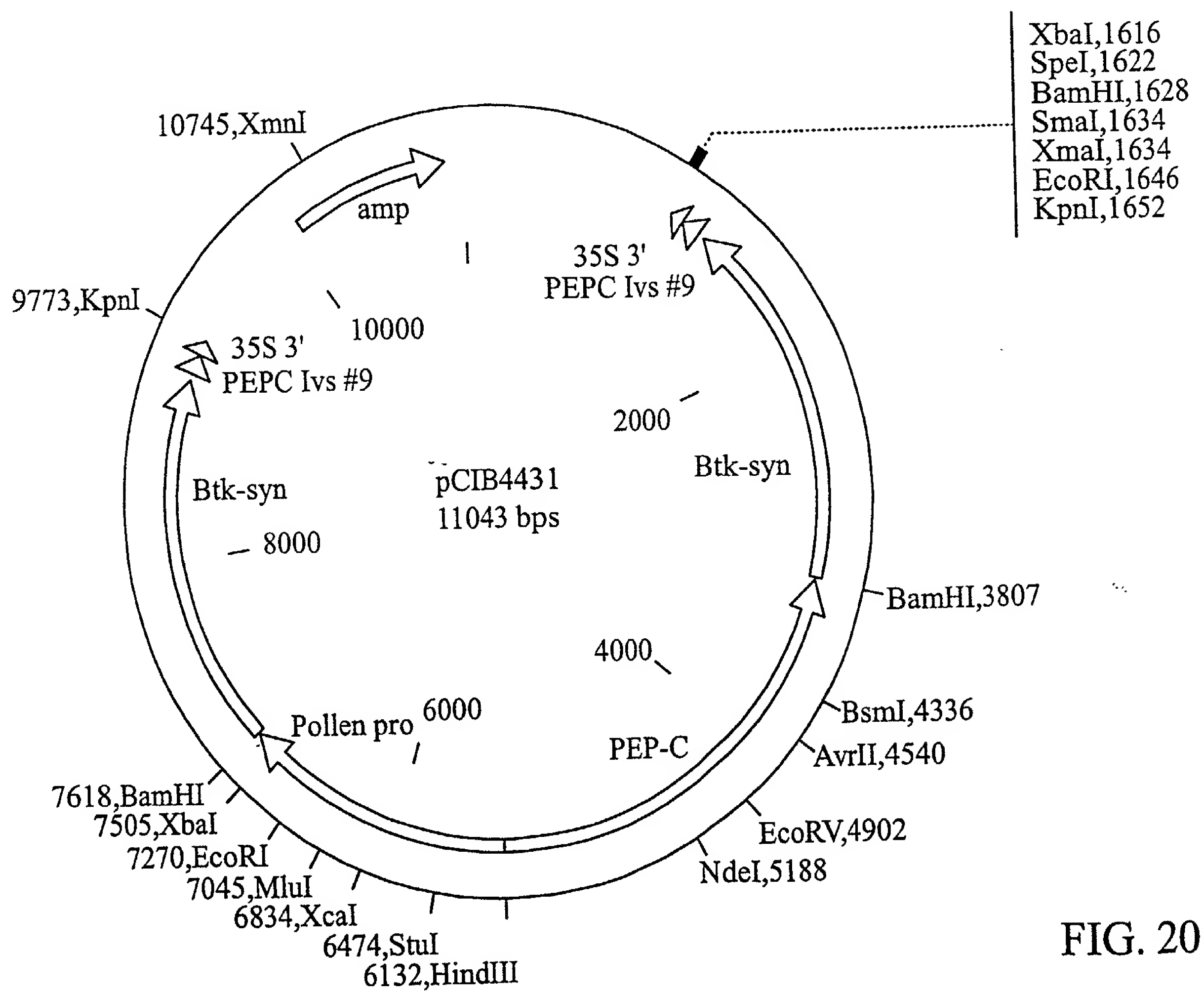

6273,BamHI
6245,NotI
6160,XbaI
5925,EcoRI
EcoNI,375
HincII,1103
SalI,1103
Btk-syn
Pollen pro
6000
5129,StuI
1000
ApaI,1470
5000
4787,HindIII
lacZ"
pCIB4429
6282 bps
2000
BstEII,1824
PEPC Ivs #9
35S 3'
SacI,1948
BglII,1955
4000
3000
KpnI,2146
EcoRI,2152
amp
NdeI,2365

FIG. 19



XbaI,1616
SpeI,1622
BamHI,1628
SmaI,1634
XmaI,1634
EcoRI,1646
KpnI,1652
10745,XmnI
amp
9773,KpnI
10000
35S 3'
PEPC Ivs #9
35S 3'
PEPC Ivs #9
2000
Btk-syn
pCIB4431
11043 bps
Btk-syn
8000
BamHI,3807
4000
Pollen pro
6000
BsmI,4336
PEP-C
AvrII,4540
7618,BamHI
7505,XbaI
7270,EcoRI
7045,MluI
6834,XcaI
6474,StuI
6132,HindIII
EcoRV,4902
NdeI,5188

FIG. 20

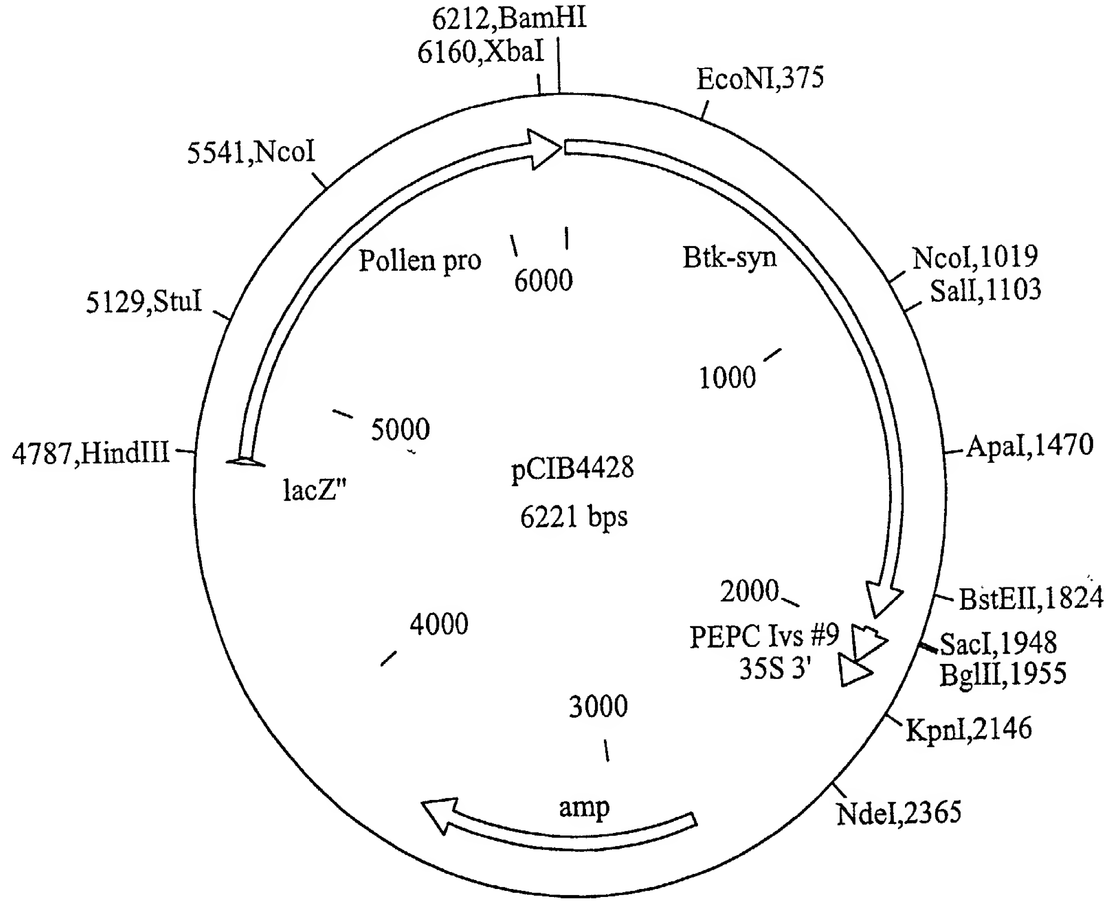

6212,BamHI
6160,XbaI
EcoNI,375
5541,NcoI
5129,StuI
4787,HindIII
NcoI,1019
SalI,1103
ApaI,1470
BstEII,1824
SacI,1948
BglII,1955
KpnI,2146
NdeI,2365
Pollen pro
6000
Btk-syn
1000
5000
lacZ"
pCIB4428
6221 bps
2000
PEPC Ivs #9
35S 3'
4000
3000
amp

FIG. 21

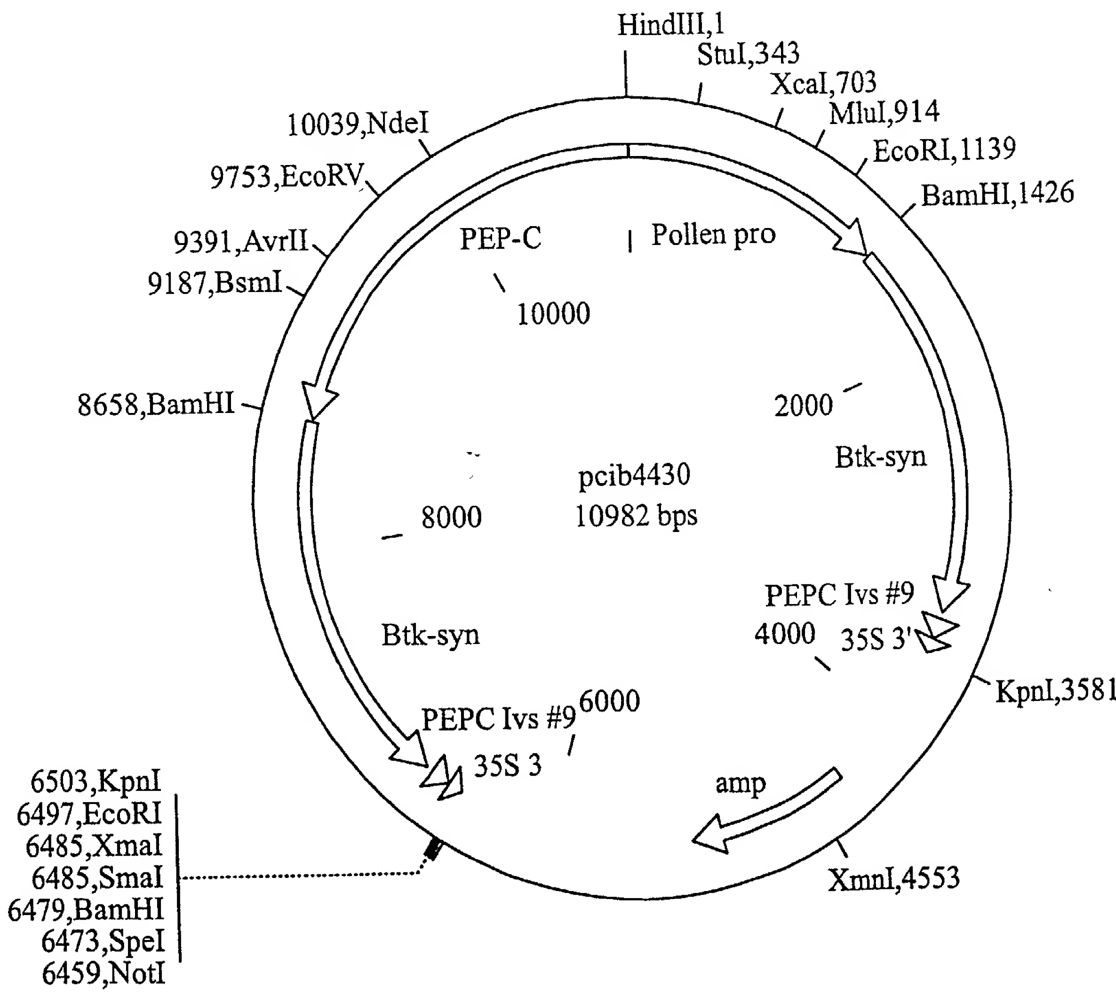

HindIII,1
StuI,343
XcaI,703
MluI,914
EcoRI,1139
BamHI,1426
KpnI,3581
XmnI,4553
6503,KpnI
6497,EcoRI
6485,XmaI
6485,SmaI
6479,BamHI
6473,SpeI
6459,NotI
8658,BamHI
9187,BsmI
9391,AvrII
9753,EcoRV
10039,NdeI
Pollen pro
2000
Btk-syn
PEPC Ivs #9
4000
35S 3'
amp
6000
PEPC Ivs #9
35S 3
Btk-syn
8000
PEP-C
10000
pcib4430
10982 bps

FIG. 22

Fig. 23A

CryIA(b) Protein Levels in Transgenic Maize

ELISA Bt Values of Field Plants:

INBRED X PARENT	ABRU PLANT Number	ng Bt/mg protein
2ND01X171-4A	1646	29
5N984X171-4A	857	1705
5N984X171-4A	870	1760
5N984X171-13	969	22
5N984X171-15	1468	17
5N984X171-15	1470	28
5N984X171-14A	1502	180
5N984X171-14A	1529	1500
5N984X176-11	1667	408
5N984X176-11	1671	1270
5N984X176-11	1673	1522
5N984X176-11	1675	943
5N984X176-11	1679	967
5N984X171-4B	1942	15
5N984X171-4B	1946	16
5NA56X171-16ABX	1101	30
5NA89X176-11	1622	959
5NA89X176-11	1630	1172
5NA89X176-11	1635	1100
6F010X171-4	825	103
6F010X171-4	832	1298

-Bt levels are in ng cryIA(b)/mg total protein.

-Data are from progeny of the described maize transformants expressing the cryIA(b) protein.

-ELISA analysis of transgenic plant material was carried out using standard procedures as described elsewhere.

Fig. 23B

Bioassay of European corn borer, Ostrinia nubilalis, and sugarcane borer, Diatraea saccharalis

Plasmid	Promoter	Cross	Plant No.	Bt Gene	Percent Mortality Ostrinia	Percent Mortality Diatraea
pCIB4431	PEPC	5N984 X 176-8B	21	+	100	100
			22	-	0	0
			40	+	100	100
pCIB4431	PEPC	5N984 X 176-11	95	+	100	100
			96	-	0	0
			98	+	100	100
pCIB4418	35S	5N984 X 171-14A	45	-	0	10
			64	+	100	90
			68	+	100	100
pCIB4431	PEPC	2N217AF X 176-8B	1	-	0	0
			3	+	100	100
			4	+	100	100
pCIB4418	35S	2N217AF X 171-15	70	-	10	0
			83	+	90	80
			88	+	90	100

Fig. 23C

CryIA(b) Protein Levels in Transgenic Maize

Greenhouse plants

35S LINE	LEAF	PITH	ROOT	POLLEN
6F010 x 171-4A	-409 + 288	NT	NT	NT
5N984 x 171-14A	256 + 159	191	198	30
6F010 x 171-16AB	240 + 174	221	271	NT
5N984 x 171-13	201 + 94	NT	NT	NT
5NA89 x 171-13	37 + 7	150	0	NT
5N984 x 171-18	7.7 + 3	NT	NT	NT
6N615 x 171-16AB	7.5 + 3	0	0	
PEPC LINE				
6N615 x 176-11	1126 + 419	41	19	NT
6F010 x 176-10	774 + 159	NT	NT	130
5N984 x 176-11	719 + 128	16	20	186

-Bt levels are in ng cryIA(b)/mg total protein.

Data are from progeny of the described maize transformants expressing the cryIA(b) protein.

ELISA analysis of transgenic plant material was carried out using standard procedures as described elsewhere.

Fig. 23D

Bioassay of European corn borer, Ostrinia nubilalis, on Pith:SynBt maize

Plasmid	Promoter	Event	Plant No.	Percent Mortality
pCIB4433	Pith	JS21A-Top	1	90
			3	80
			11	90
			13	70
			14	75
			19	85
			28	80
pCIB4433	Pith	JS22D-Mid	3	70
			4	65
			7	85
			17	95
		Control	1	5
			2	0
			3	0

Fig. 23E

EXPRESSION OF THE CRYIA(b) GENE IN TRANSGENIC MAIZE USING THE PITH-PREFERRED PROMOTER

Leaf samples from small plantlets transformed with pCIB4433 using procedures described elsewhere were analyzed for the presence of the cryIA(b) protein using ELISA. All plants expressing cryIA(b) were found to be insecticidal in the standard European corn borer bioassay.

Note that the pith-preferred promoter has a low, but detectable level of expression in leaf tissue of maize. Detection of CryIA(b) protein is consistent with this pattern of expression.

PLANT NUMBER	ng cryIA(b)/mg protein
JS21A-1 TOP	169
JS21A-2 TOP	0
JS21A-3 TOP	113
JS21A-11 TOP	127
JS21A-12 TOP	112
JS21A-13 TOP	97
JS21A-14 TOP	118
JS21A-19 TOP	82
JS21A-24 TOP	0
JS21A-28 TOP	154
JS22D-3 MID	2946
JS22D-4 MID	5590
JS22D-11 MID	215
JS22D-17 MID	3004

Fig. 24A

```
   1 GAATTCGGATCCATTAAAGAAGTCTTTGAACAGATTCTAGAGATCTAGTTTAATGAGCTC 60
  61 CCAAAAGTCTTGAAAAAATTCAGCGGGGAGGCCATTAGGGCAGGGGTACTGTTATGTTTT 120
 121 AAAGAGAACACCACTTTCTTGATCTCTTCTAAAGAGAAATGTTTTGTAAGAAGGATCCTG 180
 181 TCCTCCTCATCCAACCTTTTCATCGGCAAATTTTTCATAGAGATATTAGAGGCAAGAGAG 240
 241 GGGCCAAAAAGATCCATGTAAATGGAAGTGGCCACCTGGTTGATACCTCCCTCATCTTCA 300
 301 ACAGAAAATCCATTATGAAAAAGTGAATGGATTTTAAACTCTTCTTTTTCTTCCCTTTTG 360
 361 CAATGAGCTGAAAATATCTGGTATTATTCTCATCACCCTCATTAATGAATCTGTCCCTAG 420
 421 CAATTTGCTTTCTCTTGATCCCTTCTGCAGCCACCATGTTTCTTAAATTCCACTCCATAT 480
 481 CAAGCTTTTCCAATCTATCAGAATCTGAGATGGCTGCAATCTCTCTCATTTTCTCAAGGA 540
 541 TATCGATGTTATCCATAAGGTATTTCTTGAACTTCTTATATTTCCCTTCGACATTTATAT 600
 601 TCCATCCTTTCAACATTTTTTTGTTCAATCTTTTTTGTTTTTTTCCTTTCCAAACATCGA 660
 661 TACATTTCCTGCTCCTCACAGGTAAGGACGAGCTTTCAAAAAACCTTCTGCTTTAAAGTC 720
 721 AGGTCTGAGCCTCCAGCAAAGCTCACATATCTAAAGTCCCTCTTCTTAGTTGGGACAGAG 780
 781 TCAGTGCTAAGACACATGGGAACATGACCAGAAAAAAAAAATCATATTTAGCCCAGAGAC 840
 841 AACAATATTCTTGTACTGCAAGTCTCGTTATGGGCTAGCAAAGGAATCTACCCAACTTCT 900
 901 CAAATGTGTTGGGATGTCAAGTATATAGACTATTCATCAGTTCCAACTCTATCAAACTGT 960
 961 GCAGCTCAATTATAGAGTTGAATAAAGTGCTCCATCTATTTGTTCTTATCCTCATATTTG 1020
1021 GTTAAGATATTAAAATCACCTCCCACCAACATTTAAAGTGCACCATTTAAAGTGGCTCGC 1080
1081 GAGCACCAAACCGCTGAAAACCGGAAATGTTTAGCACGTTGGCAGCGGGACCCTTTTCTA 1140
1141 TCTCATCGTGTTCTTCGTTGTCCACCACGGCCCACGGGCCAACGCTCCTCCATCCTGTAG 1200
1201 TGTAGAGTATATTCCATTTGCGACCGAGCCGAGCATCGATCCAGCCACACTGGCCACTGC 1260
                                                 84
1261 CAGCCAGCCATGTGGCACTCCTACGTATACTACGTGAGGTGAGATTCACTCACATQGGAT 1320
-465                                                              -405

1321 GGGACCGAGATATTTTACTGCTGTGGTTGTGTGAGAGATAATAAAGCATTTATGACGATT 1380
1381 GCTGAACAGCACACACCATGCGTCCAGATAGAGAAAGCTTTCTCTCTTTATTCGCATGCA 1440

1441 TGTTTCATTATCTTTTATCATATATATATAACACATATTAAATGATTCTTCGTTCCAATT 1500
-285                                                              -226

1501 TATAATTCATTTGACTTTTTTATCCACCGATGCTCGTTTTATTAAAAAAAATATTATAAT 1560
-225                                                              -166

1561 TATTGTTACTTTTTGTTGTAATATTGTTTAGCATATAATAAACTTTGATACTAGTATGTT 1620
-165                                                              -106
                                            49
1621 TCCGAGCAAAAAAAAATATTAATATTTAGATTACGAGCCCATTAATTAATTATATTCGAG 1680
-105                                                               -46

                           83                        +1
1681 ACAAGCGAAGCAAAGCAAAGCAAGCTAATGTTGCCCCTGCTGTGCATGCAGAGGCCCGCT 1740
-45                                                                +15
                                                          73*******
1741 CTTGCTATAAACGAGGCAGCTAGACGCGACTCGACTCATCAGCCTCATCAACCTCGACGA 1800
 +16                                                               +75
     ***************
1801 AGGAGGAACGAACGGACAGGTTGTTGCACAGAAGCGACATGGCTTTCGCGCCCAAAACGT 1860
 +76                                       M   A   F   A   P   K   T   S +135
```

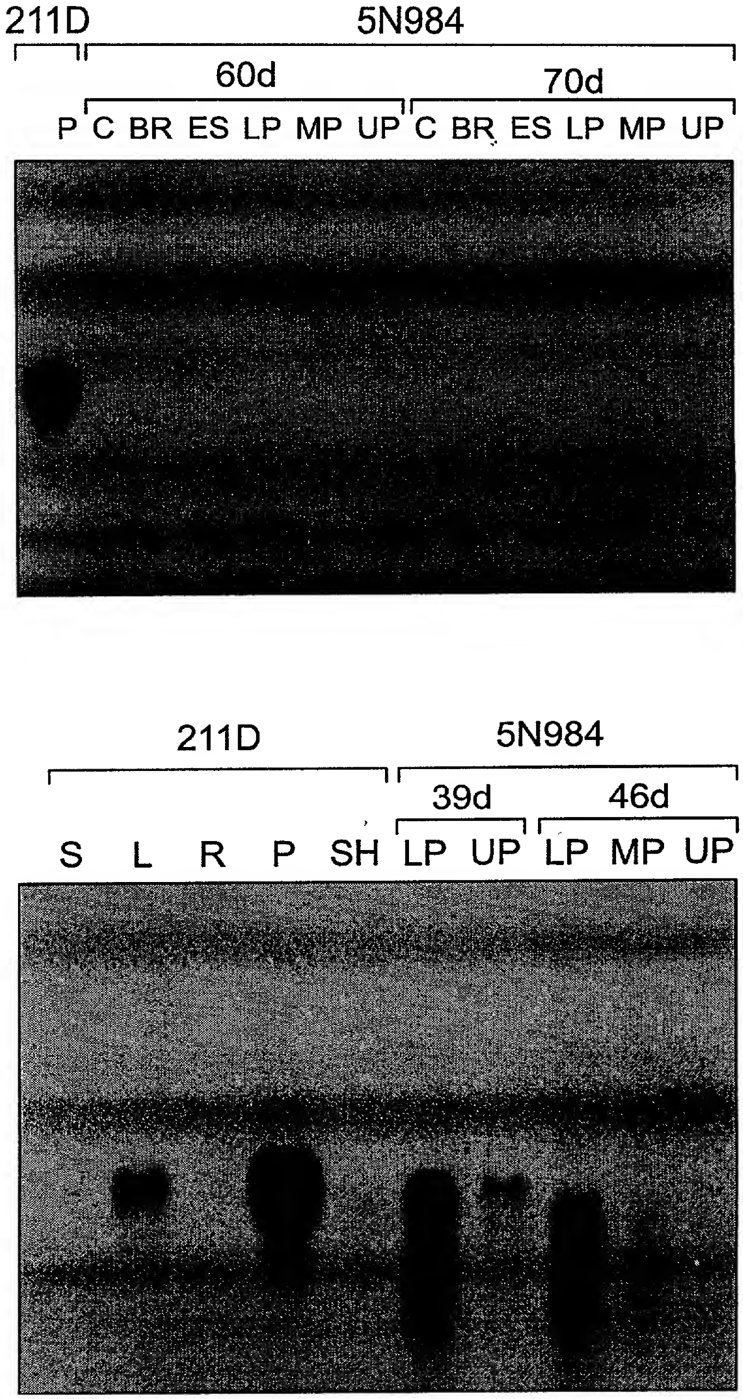
Fig. 24B

```
                                                      40 + 41
1861 CCTCCTCCTCCTCGCTGTCCTCGGCGTTGCAGGCAGCTCAGTCGCCGCCGCTGCTCCTGA 1920
+126   S  S  S  S  L  S  S  A  L  Q  A  A  Q  S  P  P  L  L  L  R +195

                                                   40 + 41
1921 GGCGGATGTCGTCGACCGCAACACCGAGACGGAGGTACGACGCGGCCGTCGTCGTCACTA 1980
+196   R  M  S  S  T  A  T  P  R  R  R  Y  D  A  A  V  V  V  T  T +255

1981 CCACCACCACTGCTAGAGCTGCGGCGGCTGCTGTCACGGTTCCCGCCGCCCCGCCGCAGG 2040
+256   T  T  T  A  R  A  A  A  A  A  V  T  V  P  A  A  P  P  Q  A +315

                                         75                    $
2041 CGGGCCGCCGCCGCCGGTGCCACCAAAGCAAGCGGCGGCACCCGCAGAGGAGGAGCCGTC 2100
+316   G  R  R  R  R  C  H  Q  S  K  R  R  H  P  Q  R  R  S  R  P +375

2101 CGGTGTCGGACACCATGGCGGCGCTCATGGCCAAGGGCAAGGTTCGTATAGTACGCGCGC 2160
+376   V  S  D  T  M  A  A  L  M  A  K  G  K

2161 GTGTCGTCGTCGTTATTTTGCGCATAGGCGCGGACATACACGTGCTTTAGCTAGCTAACA 2220
2221 GCTAGATCATCGGTGCAGACGGCGTTCATCCCGTACATCACCGCCGGCGACCCGGACCTA 2280
                    T  A  F  I  P  Y  I  T  A  G  D  P  D  L

2281 GCGACGACGGCCGAGGCGCTGCGTCTGCTGGACGGCTGTGGCGCCGACGTCATCGAGCTG 2340
     A  T  T  A  E  A  L  R  L  L  D  G  C  G  A  D  V  I  E  L

2341 GGGGTACCCTGCTCGGACCCCTACATCGACGGGCCCATCATCCAGGCGTCGGTGGCGCGG 2400
     G  V  P  C  S  D  P  Y  I  D  G  P  I  I  Q  A  S  V  A  R

2401 GCTCTGGCCAGCGGCACCACCATGGACGCCGTGCTGGAGATGCTGAGGGAGGTGACGCCG 2460
     A  L  A  S  G  T  T  M  D  A  V  L  E  M  L  R  E  V  T  P

2461 GAGCTGTCGTGCCCCGTGGTGCTCCTCTCCTACTACAAGCCCATCATGTCTCGCAGCTTG 2520
     E  L  S  C  P  V  V  L  L  S  Y  Y  K  P  I  M  S  R  S  L

2521 GCCGAGATGAAAGAGGCGGGGGTCCACGGTAACTATAGCTAGCTCTTCCGATCCCCCTTC 2580
     A  E  M  K  E  A  G  V  H

2581 AATTAATTAATTTATAGTAGTCCATTCATGTGATGATTTTTGTTTTTCTTTTTACTGACA 2640
2641 GGTCTTATAGTGCCTGATCTCCCGTACGTGGCCGCGCACTCGCTGTGGAGTGAAGCCAAG 2700
     G  L  I  V  P  D  L  P  Y  V  A  A  H  S  L  W  S  E  A  K

2701 AACAACAACCTGGAGCTGGTAGGTTGAATTAAGTTGATGCATGTGATGATTTATGTAGCT 2760
     N  N  N  L  E  L

2761 AGATCGAGCTAGCTATAATTAGGAGCATATCAGGTGCTGCTGACAACACCAGCCATACCA 2820
                                       V  L  L  T  T  P  A  I  P

2821 GAAGACAGGATGAAGGAGATCACCAAGGCTTCAGAAGGCTTCGTCTACCTGGTAGTTATA 2880
     E  D  R  M  K  E  I  T  K  A  S  E  G  F  V  Y  L

2881 TGTATATATAGATGGACGACGTAACTCATTCCAGCCCCATGCATATATGGAGGCTTCAAT 2940
2941 TCTGCAGAGACGACGAAGACCACGACGACTAACACTAGCTAGGGGCGTACGTTGCAG 3000

3001 GTGAGCGTGAACGGAGTGACAGGTCCTCGCGCAAACGTGAACCCACGAGTGGAGTCACTC 3060
     V  S  V  N  G  V  T  G  P  R  A  N  V  N  P  R  V  E  S  L
```

Fig. 24C

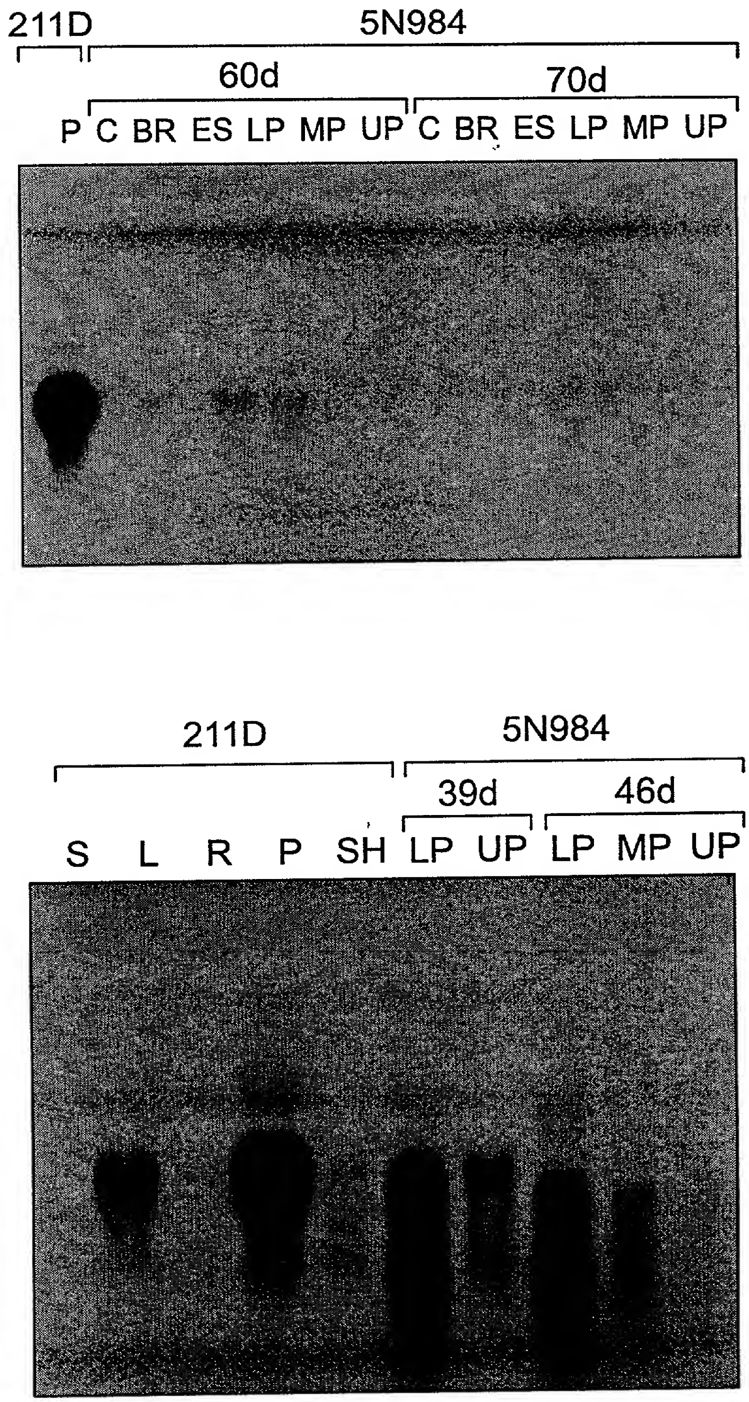
```
3061 ATCCAGGAGGTTAAGAAGGTGACTAACAAGCCCGTTGCTGTTGGCTTCGGCATATCCAAG 3120
     I  Q  E  V  K  K  V  T  N  K  P  V  A  V  G  F  G  I  S  K

3121 CCCGAGCACGTGAAGCAGGTACGTACGTAGCTGACCAAAAAAAACTGTTAACAAGTTTTG 3180
     P  E  H  V  K
3181 TTTGACAAGCCGGCTACTAGCTAGCTAACAGTGATCAGTGACACACACACACACACAGAT 3240
                                                           Q  I

3241 TGCGCAGTGGGGCGCTGACGGGGTGATCATCGGCAGCGCCATGGTGAGGCAGCTGGGCGA 3300
      A  Q  W  G  A  D  G  V  I  I  G  S  A  M  V  R  Q  L  G  E

3301 AGCGGCTTCTCCCAAGCAAGGCCTGAGGAGGCTGGAGGAGTATGCCAGGGGCATGAAGAA 3360
      A  A  S  P  K  Q  G  L  R  R  L  E  E  Y  A  R  G  M  K  N
               +++
3361 CGCGCTGCCATGAGTCCATGACAAAGTAAAACGTACAGAGACACTTGATAATATCTATCT 3420
      A  L  P
3421 ATCATCTCGGAGAAGACGACCGACCAATAAAAATAAGCCAAGTGGAAGTGAAGCTTAGCT 3480

3481 GTATATACACCGTACGTCGTCGTCGTCGTTCCGGATCGATCTCGGCCGGCTAGCTAGCAG 3540
3541 AACGTGTACGTAGTAGTATGTAATGCATGGAGTGTGGAGCTACTAGCTAGCTGGCCGTTC 3600
3601 ATTCGATTATAATTCTTCGCTCTGCTGTGGTAGCAGATGTACCTAGTCGATCTTGTACGA 3660

3661 CGAAGAAGCTGGCTAGCTAGCCGTCTCGATCGTATATGTACTGATTAATCTGCAGATTGA 3720
                                                           $
3721 ATAAAAACTACAGTACGCATATGATGCGTACGTACGTGTGTATAGTTTGTGCTCATATAT 3780
3781 GCTCCTCATCACCTGCCTGATCTGCCCATCGATCTCTCTCGTACTCCTTCCTGTTAAATG 3840
3841 CCTTCTTTGACAGACACACCACCACCAGCAGCAGTGACGCTCTGCACGCCGCCGCTTTAA 3900
3901 GACATGTAAGATATTTTAAGAGGTATAAGATACCAAGGAGCACAAATCTGGAGCACTGGG 3960
3961 ATATTGCAAAGACAAAAAAAAAACAAAATTAAAGTCCCACCAAAGTAGAGATAGTAAAGA 4020
4021 GGTGGATGGATTAAAATTATCTCATGATTTTTGGATCTGCTCAAATAGATCGATATGGTA 4080
4081 TTCAGATCTATGTTGTATAGCCTTTTCATTAGCTTTCTGAAAAAAAAATGGTATGATGAG 4140
4141 TGCGGAGTAGCTAGGGCTGTGAAGGAGTCGGATGGGCTTCCACGTACTTGTTTGTGGCCC 4200
4201 TAGTCCGGTTCTATTTAGGTCCGATCCGAGTCCGGCATGGTCCGGTTCCATACGGGCTAG 4260
4261 GACCAAGCTCGGCACGTGAGTTTTAGGCCCGTCGGCTAGCCCGAGCACGACCCGTTTTTA 4320
4321 AACTGGCTAGGACTCGCCCATTTAATAAGACAAACATTGCAAAAAATAGCTCTATTTTTT 4380
4381 ATTTAAAATATATTGTTTATTTGTGAAATGTGTATTATTTGTAATATATATTATTGTATA 4440
4441 TAGTTATATCTTCAATTATGATTTATAAATATGTTTTTTATTATGAACTCAATTTTAAGT 4500
4501 TTGATTTATGCGTTGGCGGGCTCGAGGAGGCACGGTGAACATTTTTGGGTCGGGCTTAAC 4560
4561 GGGTCGGCCCGGCCCGGTTCGGCCCATCCACGGCCCATCCCGTGTCGGCCTCGTTCGGTG 4620
4621 AGTTCAGCCCGTCGGACAACCCGTCCCCGGCCCGGATAATTAATCGGGCCTAACCGTGGC 4680
4681 GTGCTTAAACGGTCCGTGCCTCAACGGACCGGGCCGCGGGCGGCCCGTTTGACATCTCTA 4740
4741 GTGGTGTGATTAGAGATGGCGATGGGAACCGATCACTGATTCCGTGTGGAGAATTCGATA 4800
4801 TCAAGCTTATCGATACC                                            4817
```

Entire sequence of the maize TrpA gene, with introns and exons, transcription and translation strats, start and stop of cDNA.
$ = start and end of cDNA; +1 = transcription start; 73******* = primer extension primer; ▼ = start of translation; +++ = stop codon; ________ = CCAAT Box, TATAA Box, poly A addition site.
# above underlined sequences are PCR primers.



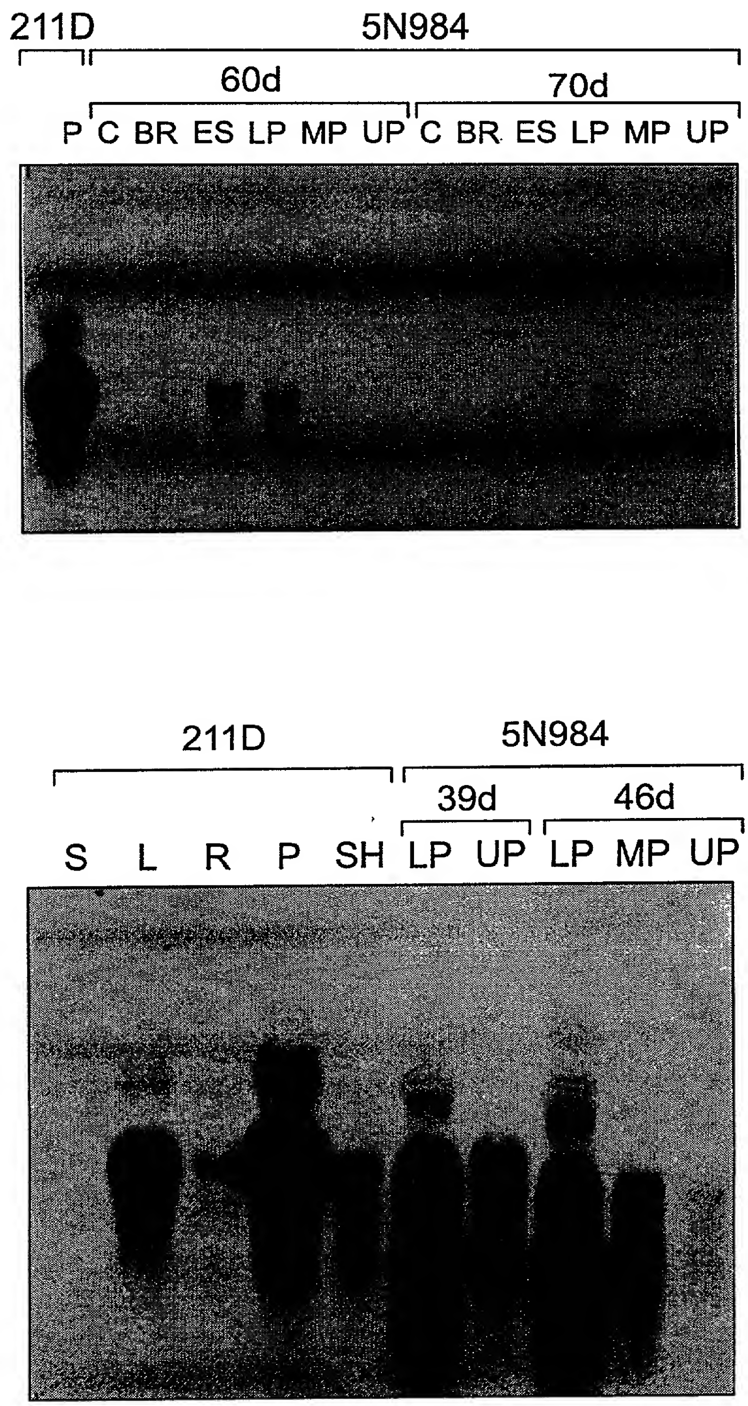
Northern blot showing differential expression of TrpA gene in maize tissues. 2 hour exposure against film at -80C with Dupont Cronex intensifying screens.

Fig. 25A



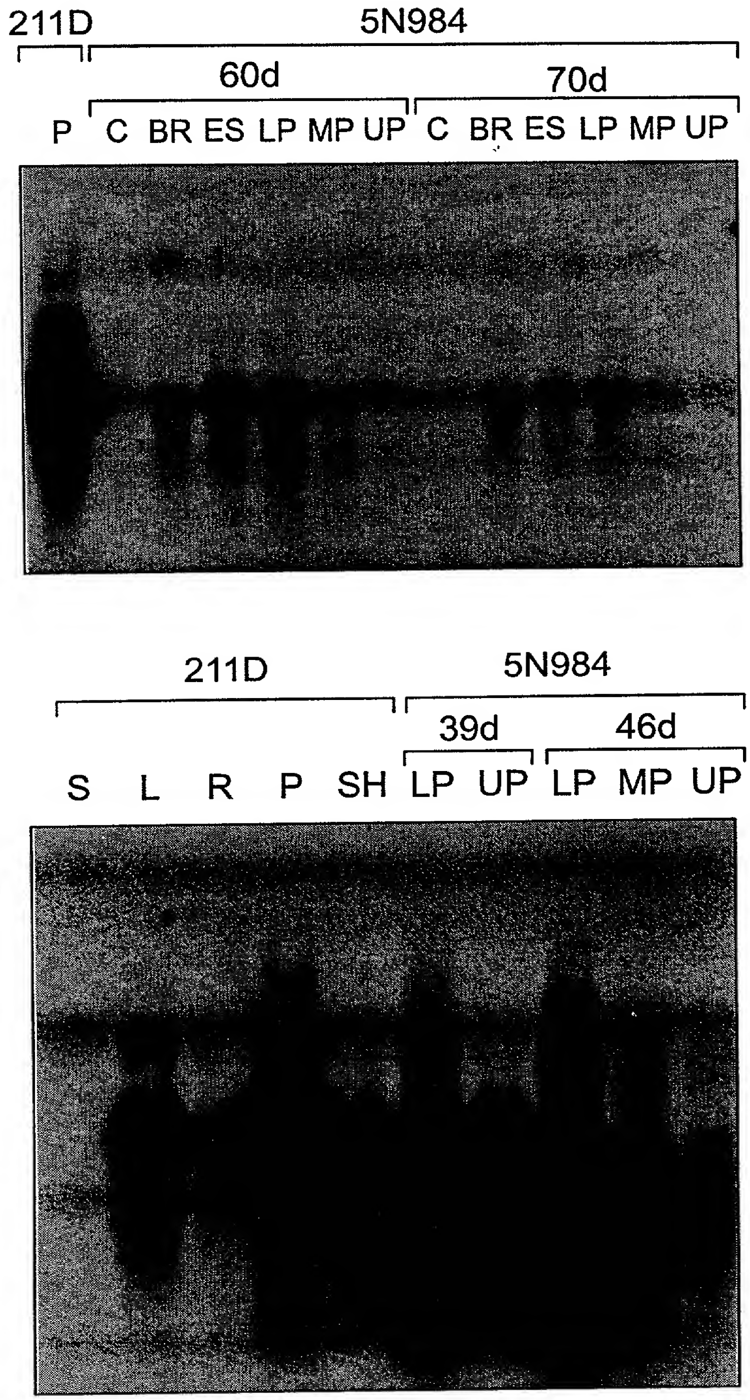
Northern blot showing differential expression of TrpA gene in maize tissues. 4 hour exposure against film at -80C with Dupont Cronex intensifying screens.

Fig. 25B



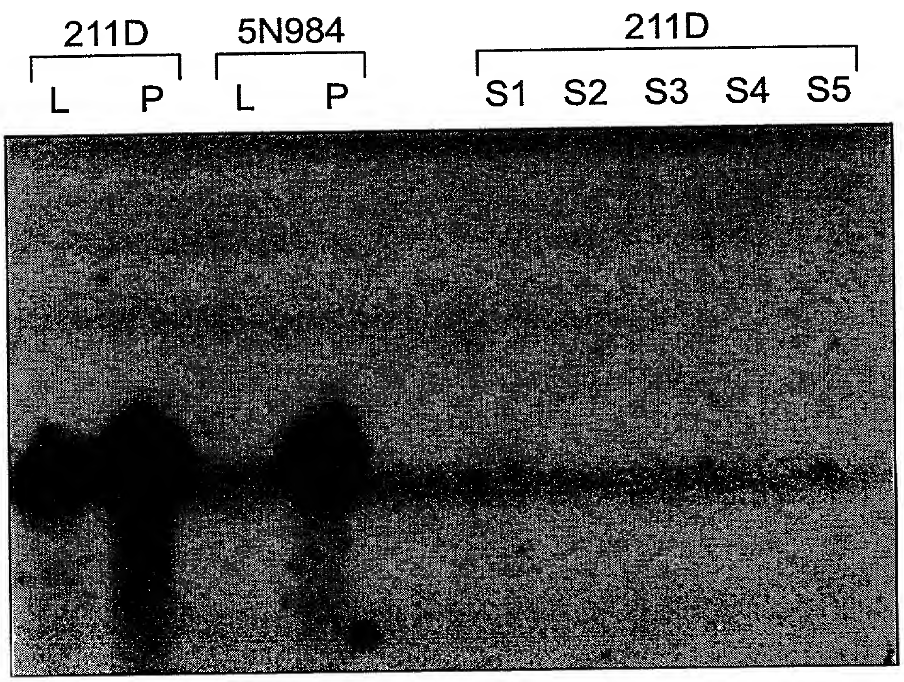
Northern blot showing differential expression of TrpA gene in maize tissues. 18 hour exposure against film at -80C with Dupont Cronex intensifying screens.

Fig. 25C



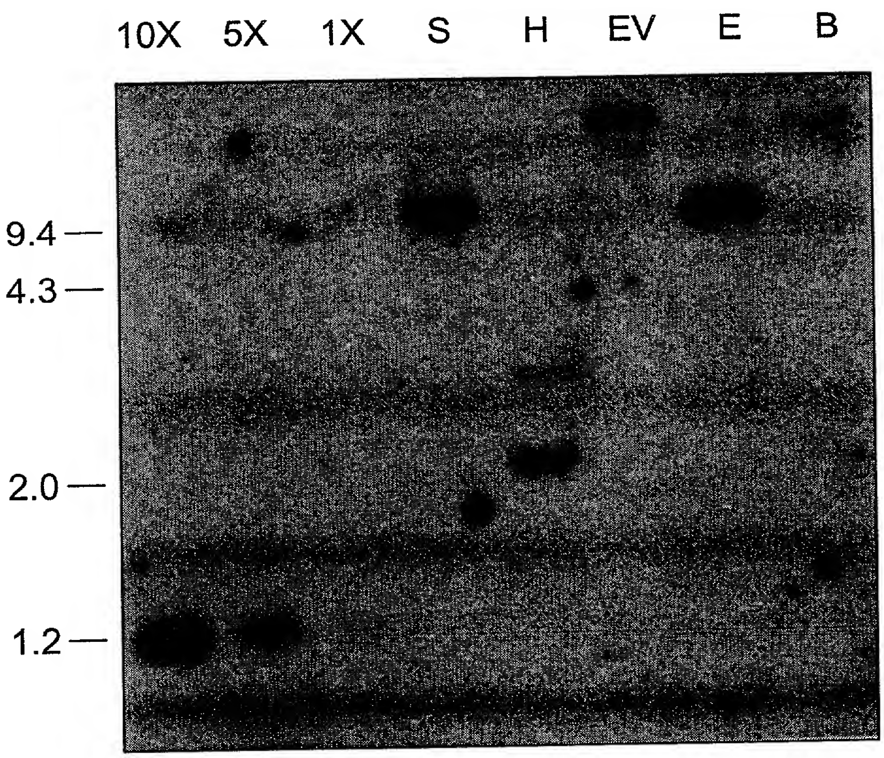
Northern blot showing differential expression of TrpA gene in maize tissues. 48 hour exposure against film at -80C with Dupont Cronex intensifying screens.

Fig. 25D



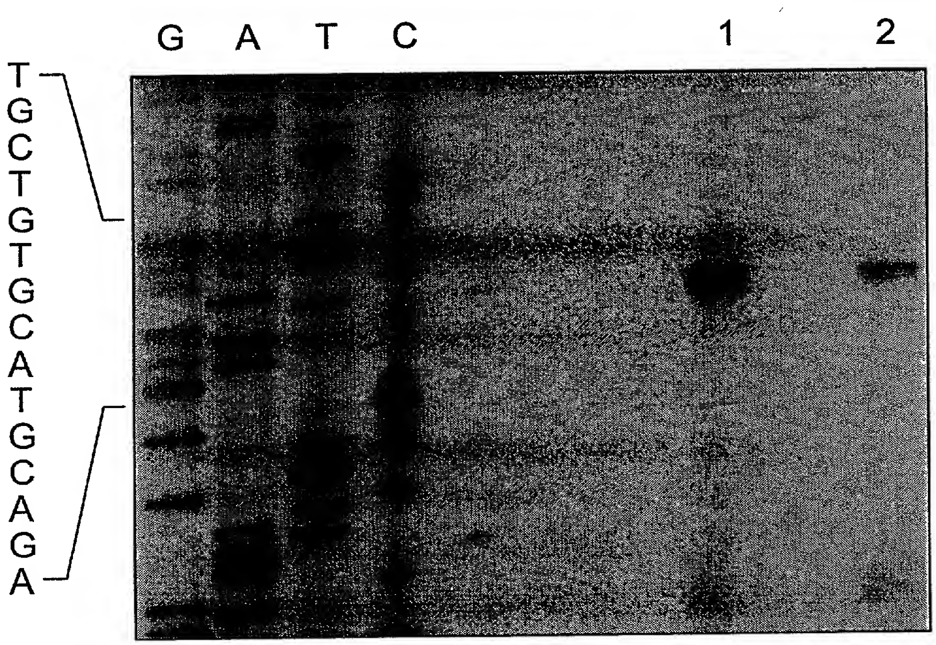
Northern blot showing maize TrpA gene expression in Funk lines 211D and 5N984 leaf and pith and the absense of expression in 211D seed total RNA.
65 hour exposure against film at -80C with Dupont Cronex intensifying screens.

Fig. 26



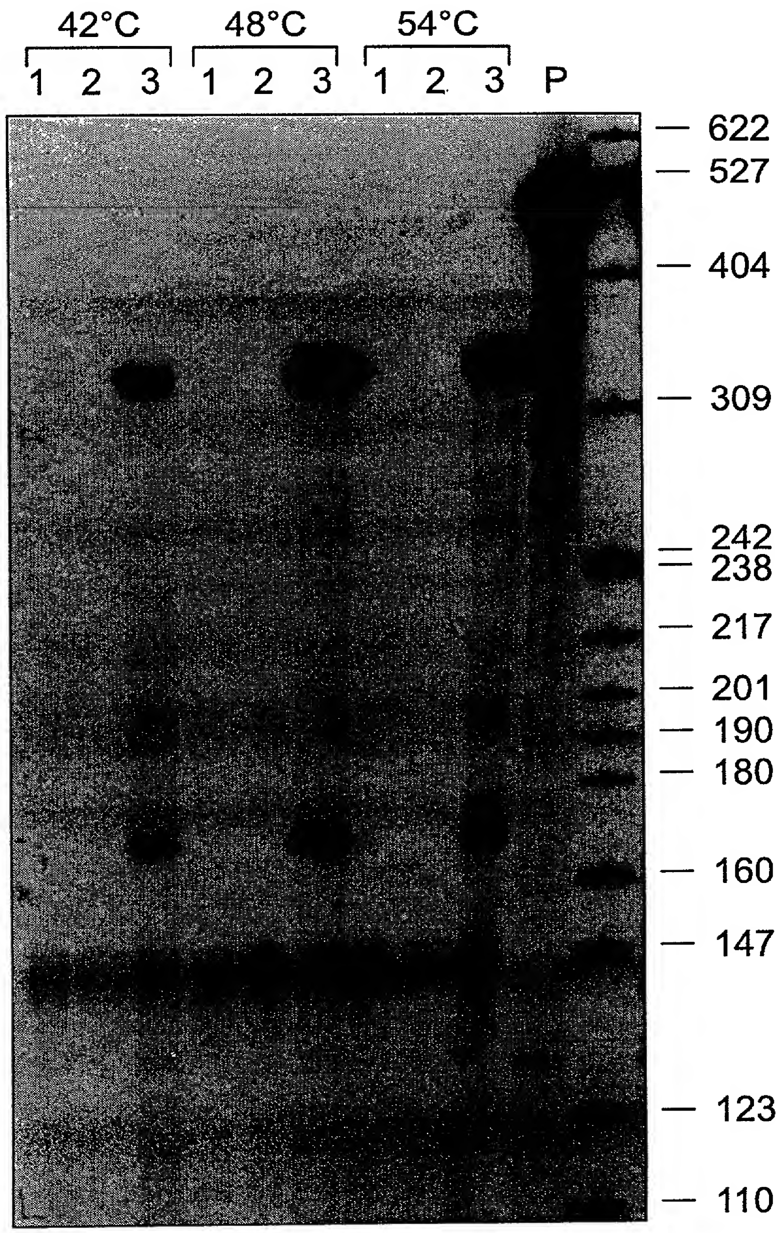
Genomic southern of Funk line 211D probed with the TrpA cDNA 8-2. B = BamHI, E = EcoRI, EV = EcoRV, H = HindIII and S = SacI.
120 hour exposure against film at -80C with Dupont Cronex intensifying screens.

Fig. 27



Primer extension showing the transcription start of TrpA gene and sequencing ladder.
1 hour exposure against film at -80C with Dupont Cronex intensifying screens.

Fig. 28A



RNase protection of region from +2 bp to +387 bp
with three annealing temperatures.
16 hour exposure against film at -80C with Dupont Cronex
intensifying screens.

Fig. 28B

Fig. 30B

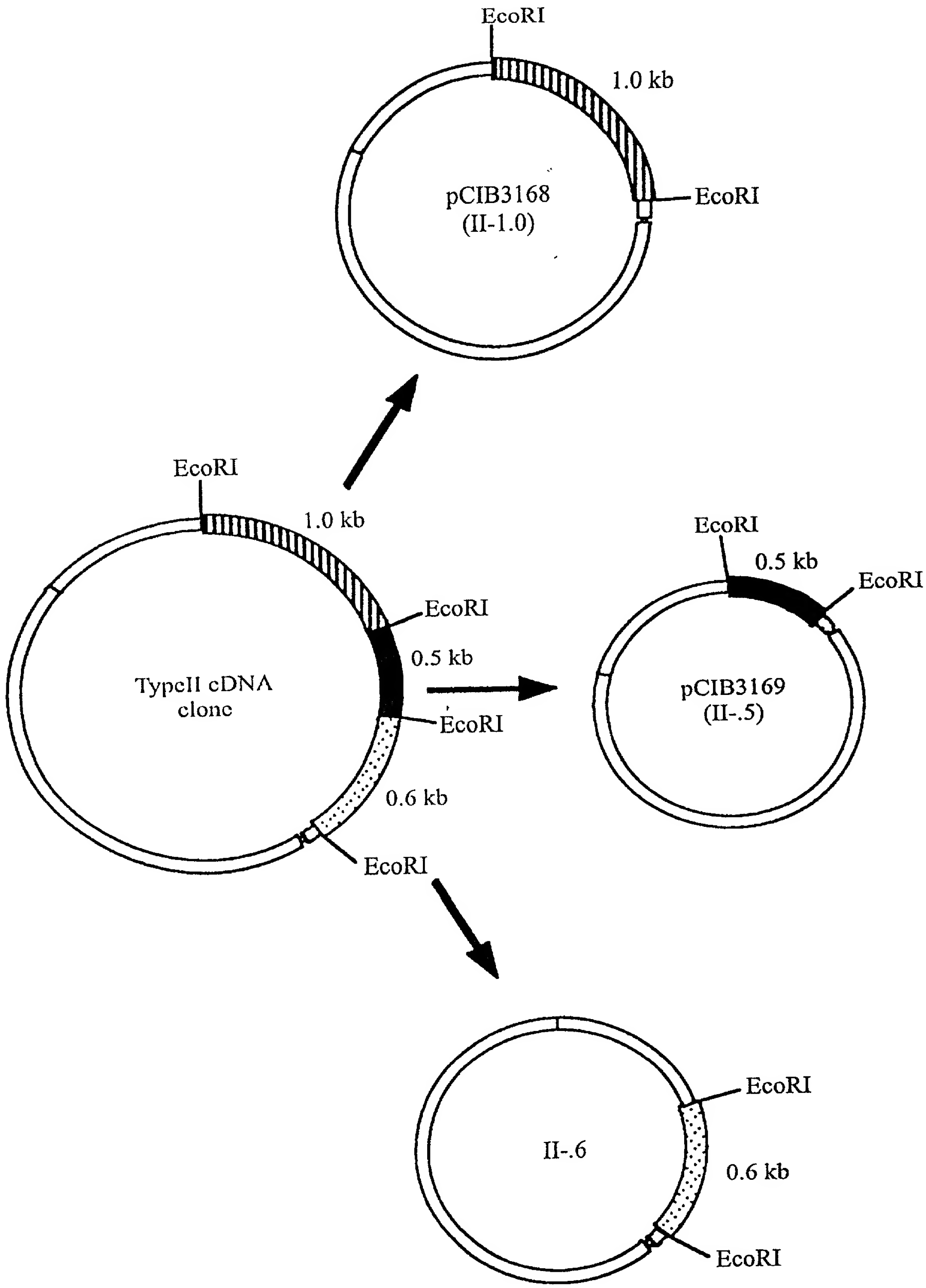
```
969 CTG GAT AGA GAA GAG CAC CTT TAC ACA GCA TTC CAG TAT TTC GAC AAG GAC AAC AGC
323▶Leu Asp Arg Glu Glu His Leu Tyr Thr Ala Phe Gln Tyr Phe Asp Lys Asp Asn Ser

1026 GGG TAC ATT ACT AAA GAA GAG CTT GAG CAC GCC TTG AAG GAG CAA GGG TTG TAT GAC
342▶Gly Tyr Ile Thr Lys Glu Glu Leu Glu His Ala Leu Lys Glu Gln Gly Leu Tyr Asp

1083 GCC GAT AAA ATC AAA GAC ATC ATC TCC GAT GCC GAC TCT GAC AAT GAT GGA AGG ATA
361▶Ala Asp Lys Ile Lys Asp Ile Ile Ser Asp Ala Asp Ser Asp Asn Asp Gly Arg Ile

1140 GAT TAT TCA GAG TTT GTG GCG ATG ATG AGG AAA GGG ACG GCT GGT GCC GAG CCA ATG
380▶Asp Tyr Ser Glu Phe Val Ala Met Met Arg Lys Gly Thr Ala Gly Ala Glu Pro Met

1197 AAC ATC AAG AAG AGG CGA GAC ATA GTC CTA TAG TGAAGTGAAGCAGCAAGTGTGTAATGTAATGTG
399▶Asn Ile Lys Lys Arg Arg Asp Ile Val Leu ...

1263 TATAGCAGCTCAAACAAGCAAATTTGTACATCTGTACACAAATGCAATGGGGTTACTTTTGCAAAAAAAAAAAAAAA
                j
1340 AAAAAAAAAA
```

Fig. 30A

Maize Pollen CDPK cDNA sequence
sequence contained in clones pCIB3168 and pCIB3169

```
  1 TG CAG ATC ATG CAC CAC CTC TCC GGC CAG CCC AAC GTG GTG GGC CTC CGC GGC GCG
    1▶Gln Ile Met His His Leu Ser Gly Gln Pro Asn Val Val Gly Leu Arg Gly Ala

 57 TAC GAG GAC AAG CAG AGC GTG CAC CTC GTC ATG GAG CTG TGC GCG GGC GGG GAG CTC
 19▶Tyr Glu Asp Lys Gln Ser Val His Leu Val Met Glu Leu Cys Ala Gly Gly Glu Leu

                              AvaI
114 TTC GAC CGC ATC ATC GCC CGG GGC CAG TAC ACG GAG CGC GGC GCC GCG GAG CTG CTG
 38▶Phe Asp Arg Ile Ile Ala Arg Gly Gln Tyr Thr Glu Arg Gly Ala Ala Glu Leu Leu

171 CGC GCC ATC GTG CAG ATC GTG CAC ACC TGC CAC TCC ATG GGG GTG ATG CAC CGG GAC
 57▶Arg Ala Ile Val Gln Ile Val His Thr Cys His Ser Met Gly Val Met His Arg Asp

           AvaI
228 ATC AAG CCC GAG AAC TTC CTG CTG CTC AGC AAG GAC GAG GAC GCG CCG CTC AAG GCC
 76▶Ile Lys Pro Glu Asn Phe Leu Leu Leu Ser Lys Asp Glu Asp Ala Pro Leu Lys Ala

285 ACC GAC TTC GGC CTC TCC GTC TTC TTC AAG GAG GGC GAG CTG CTC AGG GAC ATC GTC
 95▶Thr Asp Phe Gly Leu Ser Val Phe Phe Lys Glu Gly Glu Leu Leu Arg Asp Ile Val

                                   AvaI
342 GGC AGC GCC TAC TAC ATC GCG CCC GAG GTG CTC AAG AGG AAG TAC GGC CCG GAG GCC
114▶Gly Ser Ala Tyr Tyr Ile Ala Pro Glu Val Leu Lys Arg Lys Tyr Gly Pro Glu Ala

399 GAC ATC TGG AGC GTC GGC GTC ATG CTC TAC ATC TTC CTC GCC GGC GTG CCT CCC TTC
133▶Asp Ile Trp Ser Val Gly Val Met Leu Tyr Ile Phe Leu Ala Gly Val Pro Pro Phe

456 TGG GCA GAG AAC GAG AAC GGC ATC TTC ACC GCC ATC CTG CGA GGG CAG CTT GAC CTC
152▶Trp Ala Glu Asn Glu Asn Gly Ile Phe Thr Ala Ile Leu Arg Gly Gln Leu Asp Leu

513 TCC AGC GAG CCA TGG CCA CAC ATC TCG CCG GGA GCC AAG GAT CTC GTC AAG AAG ATG
171▶Ser Ser Glu Pro Trp Pro His Ile Ser Pro Gly Ala Lys Asp Leu Val Lys Lys Met

570 CTC AAC ATC AAC CCC AAG GAG CGG CTC ACG GCG TTC CAG GTC CTC AAT CAC CCA TGG
190▶Leu Asn Ile Asn Pro Lys Glu Arg Leu Thr Ala Phe Gln Val Leu Asn His Pro Trp

627 ATC AAA GAA GAC GGA GAC GCG CCT GAC ACG CCG CTT GAC AAC GTT GTT CTC GAC AGG
209▶Ile Lys Glu Asp Gly Asp Ala Pro Asp Thr Pro Leu Asp Asn Val Val Leu Asp Arg

684 CTC AAG CAG TTC AGG GCC ATG AAC CAG TTC AAG AAA GCA GCA TTG AGG ATC ATA GCT
228▶Leu Lys Gln Phe Arg Ala Met Asn Gln Phe Lys Lys Ala Ala Leu Arg Ile Ile Ala

741 GGG TGC CTA TCC GAA GAG GAG ATC ACA GGG CTG AAG GAG ATG TTC AAG AAC ATT GAC
247▶Gly Cys Leu Ser Glu Glu Glu Ile Thr Gly Leu Lys Glu Met Phe Lys Asn Ile Asp

798 AAG GAT AAC AGC GGG ACC ATT ACC CTC GAC GAG CTC AAA CAC GGG TTG GCA AAG CAC
266▶Lys Asp Asn Ser Gly Thr Ile Thr Leu Asp Glu Leu Lys His Gly Leu Ala Lys His

855 GGG CCC AAG CTG TCA GAC AGC GAA ATG GAG AAA CTA ATG GAA GCA GCT GAC GCT GAC
285▶Gly Pro Lys Leu Ser Asp Ser Glu Met Glu Lys Leu Met Glu Ala Ala Asp Ala Asp

                                          EcoRI
912 GGC AAC GGG TTA ATT GAC TAC GAC GAA TTC GTC ACC GCA ACA GTG CAT ATG AAC AAA
304▶Gly Asn Gly Leu Ile Asp Tyr Asp Glu Phe Val Thr Ala Thr Val His Met Asn Lys
```

Fig. 29


EcoRI
1.0 kb
pCIB3168
(II-1.0)
EcoRI
EcoRI
1.0 kb
EcoRI
0.5 kb
TypeII cDNA
clone
EcoRI
0.6 kb
EcoRI
EcoRI
0.5 kb
EcoRI
pCIB3169
(II-.5)
EcoRI
II-.6
0.6 kb
EcoRI

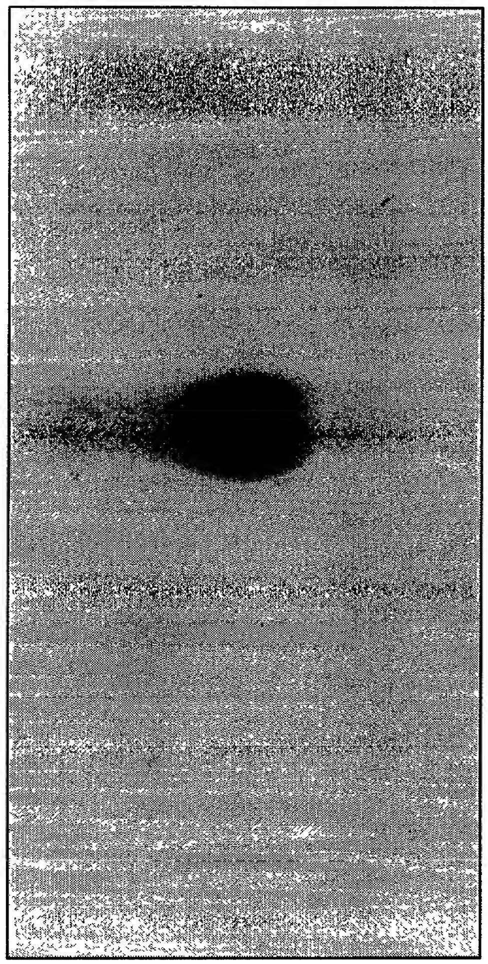

inner leaf sheath
inner leaf whorl
green leaf
anther
pollen
silk
kernel
pith
root

Fig. 31

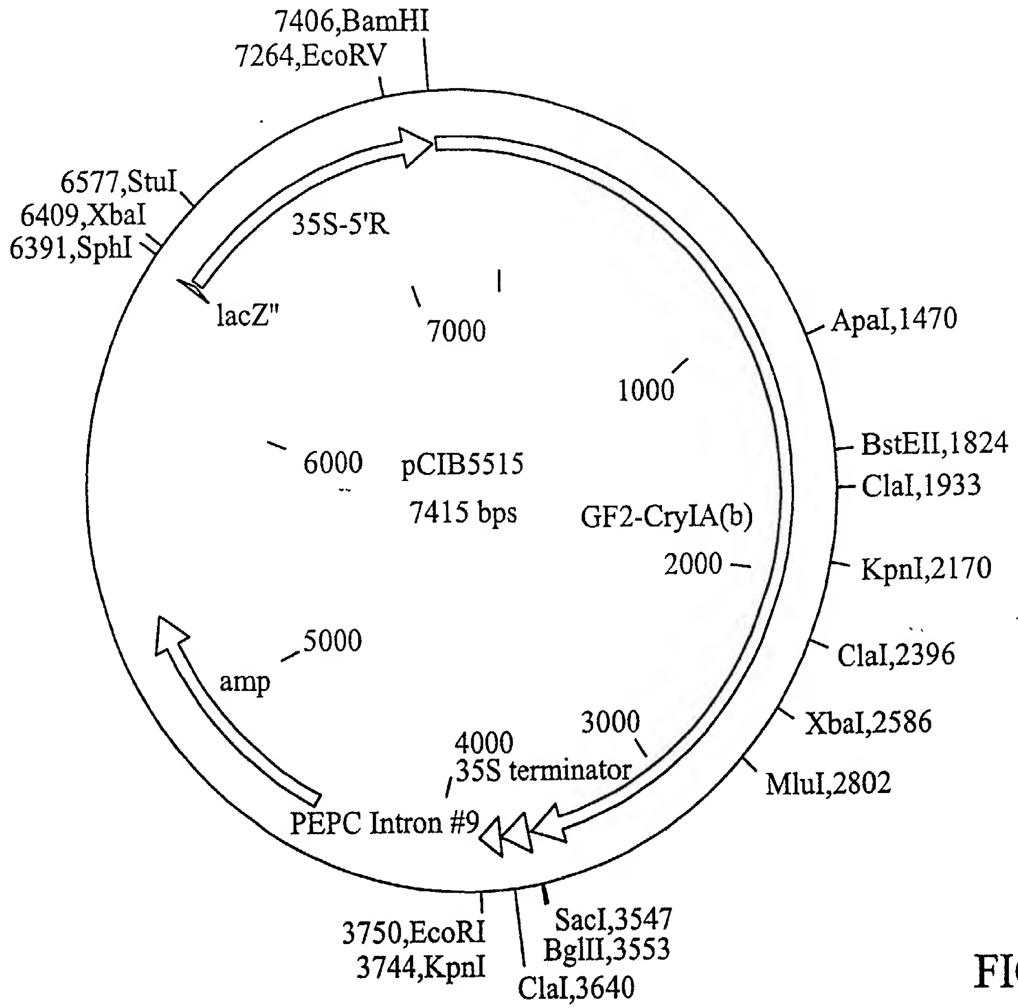

7406,BamHI
7264,EcoRV
6577,StuI
6409,XbaI
6391,SphI
35S-5'R
lacZ"
7000
1000
ApaI,1470
6000
pCIB5515
7415 bps
GF2-CryIA(b)
BstEII,1824
ClaI,1933
2000
KpnI,2170
5000
amp
ClaI,2396
XbaI,2586
3000
4000
35S terminator
MluI,2802
PEPC Intron #9
3750,EcoRI
3744,KpnI
SacI,3547
BglII,3553
ClaI,3640

FIG. 38

Fig. 37C

```
2641 AGAGCGGAGA 2VGWkTGGAG AGACAAACGT GAAAAATTGG AATGGGAAAC AAATATTGTT
     ArgAlaGlu LysLysTrp ArgAspLysArg GluLysLeu GluTrpGlu ThrAsnIleVal

2701 TATAAAGAGG CAAAAGAATC TGTAGATGCT TTATTTGTAA ACTCTCAATA TGATAGATTA
     TyrLysGlu AlaLysGlu SerValAspAla LeuPheVal AsnSerGln TyrAspArgLeu

2761 CAAGCGGATA CCAACATCGC GATGATTCAT GCGGCAGATA AACGCGTTCA TAGCATTCGA
     GlnAlaAsp ThrAsnIle AlaMetIleHis AlaAlaAsp LysArgVal HisSerIleArg

2821 GAAGCTTATC TGCCTGAGCT GTCTGTGATT CCGGGTGTCA ATGCGGCTAT TTTTGAAGAA
     GluAlaTyr LeuProGlu LeuSerValIle ProGlyVal AsnAlaAla IlePheGluGlu

2881 TTAGAAGGGC GTATTTTCAC TGCATTCTCC CTATATGATG CGAGAAATGT CATTAAAAAT
     LeuGluGly ArgIlePhe ThrAlaPheSer LeuTyrAsp AlaArgAsn ValIleLysAsn

2941 GGTGATTTTA ATAATGGCTT ATCCTGCTGG AACGTGAAAG GGCATGTAGA TGTAGAAGAA
     GlyAspPhe AsnAsnGly LeuSerCysTrp AsnValLys GlyHisVal AspValGluGlu

3001 CAAAACAACC ACCGTTCGGT CCTTGTTGTT CCGGAATGGG AAGCAGAAGT GTCACAAGAA
     GlnAsnAsn HisArgSer ValLeuValVal ProGluTrp GluAlaGlu ValSerGlnGlu

3061 GTTCGTGTCT GTCCGGGTCG TGGCTATATC CTTCGTGTCA CAGCGTACAA GGAGGGATAT
     ValArgVal CysProGly ArgGlyTyrIle LeuArgVal ThrAlaTyr LysGluGlyTyr

3121 GGAGAAGGTT GCGTAACCAT TCATGAGATC GAGAACAATA CAGACGAACT GAAGTTTAGC
     GlyGluGly CysValThr IleHisGluIle GluAsnAsn ThrAspGlu LeuLysPheSer

3181 AACTGTGTAG AAGAGGAAGT ATATCCAAAC AACACGGTAA CGTGTAATGA TTATACTGCG
     AsnCysVal GluGluGlu ValTyrProAsn AsnThrVal ThrCysAsn AspTyrThrAla

3241 ACTCAAGAAG AATATGAGGG TACGTACACT TCTCGTAATC GAGGATATGA CGGAGCCTAT
     ThrGlnGlu GluTyrGlu GlyThrTyrThr SerArgAsn ArgGlyTyr AspGlyAlaTyr

3301 GAAAGCAATT CTTCTGTACC AGCTGATTAT GCATCAGCCT ATGAAGAAAA AGCATATACA
     GluSerAsn SerSerVal ProAlaAspTyr AlaSerAla TyrGluGlu LysAlaTyrThr

3361 GATGGACGAA GAGACAATCC TTGTGAATCT AACAGAGGAT ATGGGGATTA CACACCACTA
     AspGlyArg ArgAspAsn ProCysGluSer AsnArgGly TyrGlyAsp TyrThrProLeu

3421 CCAGCTGGCT ATGTGACAAA AGAATTAGAG TACTTCCCAG AAACCGATAA GGTATGGATT
     ProAlaGly TyrValThr LysGluLeuGlu TyrPhePro GluThrAsp LysValTrpIle

3481 GAGATCGGAG AAACGGAAGG AACATTCATC GTGGACAGCG TGGAATTACT TCTTATGGAG
     GluIleGly GluThrGlu GlyThrPheIle ValAspSer ValGluLeu LeuLeuMetGlu

3541 GAATAA
     Glu---
```

Fig. 37B

```
1261 CCACCTCGAC AGGGCTTCAG CCACCGTCTG AGCCACGTGA CCATGTTCCG CAGTGGCTTC
     ProProArg GlnGlyPhe SerHisArgLeu SerHisVal SerMetPhe ArgSerGlyPhe

1321 AGCAACAGCA GCGTGAGCAT CATCCGTGCA CCTATGTTCA GCTGGATTCA CCGCAGTGCC
     SerAsnSer SerValSer IleIleArgAla ProMetPhe SerTrpIle HisArgSerAla

1381 GAGTTCAACA ACATCATCCC CAGCAGCCAG ATCACCCAGA TCCCCCTGAC CAAGAGCACC
     GluPheAsn AsnIleIle ProSerSerGln IleThrGln IleProLeu ThrLysSerThr

1441 AACCTGGGCA GCGGCACCAG CGTGGTGAAG GGCCCCGGCT TCACCGGCGG CGACATCCTG
     AsnLeuGly SerGlyThr SerValValLys GlyProGly PheThrGly GlyAspIleLeu

1501 CGCCGCACCA GCCCCGGCCA GATCAGCACC CTGCGCGTGA ACATCACCGC CCCCCTGAGC
     ArgArgThr SerProGly GlnIleSerThr LeuArgVal AsnIleThr AlaProLeuSer

1561 CAGCGCTACC GCGTCCGCAT CCGCTACGCC AGCACCACCA ACCTGCAGTT CCACACCAGC
     GlnArgTyr ArgValArg IleArgTyrAla SerThrThr AsnLeuGln PheHisThrSer

1621 ATCGACGGCC GCCCCATCAA CCAGGGCAAC TTCAGCGCCA CCATGAGCAG CGGCAGCAAC
     IleAspGly ArgProIle AsnGlnGlyAsn PheSerAla ThrMetSer SerGlySerAsn

1681 CTGCAGAGCG GCAGCTTCCG CACCGTGGGC TTCACCACCC CCTTCAACTT CAGCAACGGC
     LeuGlnSer GlySerPhe ArgThrValGly PheThrThr ProPheAsn PheSerAsnGly

1741 AGCAGCGTGT TCACCCTGAG CGCCCACGTG TTCAACAGCG GCAACGAGGT GTACATCGAC
     SerSerVal PheThrLeu SerAlaHisVal PheAsnSer GlyAsnGlu ValTyrIleAsp

1801 CGCATCGAGT TCGTGCCCGC CGAGGTGACC TTCGAGGCCG AGTACGACCT GGAGAGGGCT
     ArgIleGlu PheValPro AlaGluValThr PheGluAla GluTyrAsp LeuGluArgAla

1861 CAGAAGGCCG TGAACGAGCT GTTCACCAGC AGCAACCAGA TCGGCCTGAA GACCGACGTG
     GlnLysAla ValAsnGlu LeuPheThrSer SerAsnGln IleGlyLeu LysThrAspVal

1921 ACCGACTACC ACATCGATCA AGTATCCAAT TTAGTTGAGT GTTTATCTGA TGAATTTTGT
     ThrAspTyr HisIleAsp GlnValSerAsn LeuValGlu CysLeuSer AspGluPheCys

1981 CTGGATGAAA AAAAAGAATT GTCCGAGAAA GTCAAACATG CGAAGCGACT TAGTGATGAG
     LeuAspGlu LysLysGlu LeuSerGluLys ValLysHis AlaLysArg LeuSerAspGlu

2041 CGGAATTTAC TTCAAGATCC AAACTTTAGA GGGATCAATA GACAACTAGA CCGTGGCTGG
     ArgAsnLeu LeuGlnAsp ProAsnPheArg GlyIleAsn ArgGlnLeu AspArgGlyTrp

2101 AGAGGAAGTA CGGATATTAC CATCCAAGGA GGCGATGACG TATTCAAAGA GAATTACGTT
     ArgGlySer ThrAspIle ThrIleGlnGly GlyAspAsp ValPheLys GluAsnTyrVal

2161 ACGCTATTGG GTACCTTTGA TGAGTGCTAT CCAACGTATT TATATCAAAA AATAGATGAG
     ThrLeuLeu GlyThrPhe AspGluCysTyr ProThrTyr LeuTyrGln LysIleAspGlu

2221 TCGAAATTAA AAGCCTATAC CCGTTACCAA TTAAGAGGGT ATATCGAAGA TAGTCAAGAC
     SerLysLeu LysAlaTyr ThrArgTyrGln LeuArgGly TyrIleGlu AspSerGlnAsp

2281 TTAGAAATCT ATTTAATTCG CTACAATGCC AAACACGAAA CAGTAAATGT GCCAGGTACG
     LeuGluIle TyrLeuIle ArgTyrAsnAla LysHisGlu ThrValAsn ValProGlyThr

2341 GGTTCCTTAT GGCCGCTTTC AGCCCCAAGT CCAATCGGAA AATGTGGGGA GCCGAATCGA
     GlySerLeu TrpProLeu SerAlaProSer ProIleGly LysCysGly GluProAsnArg

2401 TGCGCTCCGC ACCTGGAGTG GAACCCGGAC CTAGACTGCA GCTGCAGGGA CGGGGAGAAG
     CysAlaPro HisLeuGlu TrpAsnProAsp LeuAspCys SerCysArg AspGlyGluLys

2461 TGCGCCCATC ATTCCCATCA TTTCTCCTTG GACATTGATG TTGGATGTAC AGACTTAAAT
     CysAlaHis HisSerHis HisPheSerLeu AspIleAsp ValGlyCys ThrAspLeuAsn

2521 GAGGACTTAG GTGTATGGGT GATATTCAAG ATTAAGACGC AAGATGGCCA TGCAAGACTA
     GluAspLeu GlyValTrp ValIlePheLys IleLysThr GlnAspGly HisAlaArgLeu

2581 GGAAATCTAG AATTTCTCGA AGAGAAACCA TTAGTAGGAG AAGCACTAGC TCGTGTGAAA
     GlyAsnLeu GluPheLeu GluGluLysPro LeuValGly GluAlaLeu AlaArgValLys
```

Fig. 37A

```
   1 ATGGACAACA ACCCCAACAT CAACGAGTGC ATCCCCTACA ACTGCCTGAG CAACCCCGAG
     MetAspAsn AsnProAsn IleAsnGluCys IleProTyr AsnCysLeu SerAsnProGlu

  61 GTGGAGGTGC TGGGCGGCGA GCGCATCGAG ACCGGCTACA CCCCCATCGA CATCAGCCTG
     ValGluVal LeuGlyGly GluArgIleGlu ThrGlyTyr ThrProIle AspIleSerLeu

 121 AGCCTGACCC AGTTCCTGCT GAGCGAGTTC GTGCCCGGCG CCGGCTTCGT GCTGGGCCTG
     SerLeuThr GlnPheLeu LeuSerGluPhe ValProGly AlaGlyPhe ValLeuGlyLeu

 181 GTGGACATCA TCTGGGGCAT CTTCGGCCCC AGCCAGTGGG ACGCCTTCCT GGTGCAGATC
     ValAspIle IleTrpGly IlePheGlyPro SerGlnTrp AspAlaPhe LeuValGlnIle

 241 GAGCAGCTGA TCAACCAGCG CATCGAGGAG TTCGCCCGCA ACCAGGCCAT CAGCCGCCTG
     GluGlnLeu IleAsnGln ArgIleGluGlu PheAlaArg AsnGlnAla IleSerArgLeu

 301 GAGGGCCTGA GCAACCTGTA CCAAATCTAC GCCGAGAGCT TCCGCGAGTG GGAGGCCGAC
     GluGlyLeu SerAsnLeu TyrGlnIleTyr AlaGluSer PheArgGlu TrpGluAlaAsp

 361 CCCACCAACC CCGCCCTGCG CGAGGAGATG CGCATCCAGT TCAACGACAT GAACAGCGCC
     ProThrAsn ProAlaLeu ArgGluGluMet ArgIleGln PheAsnAsp MetAsnSerAla

 421 CTGACCACCG CCATCCCCCT GTTCGCCGTG CAGAACTACC AGGTGCCCCT GCTGAGCGTG
     LeuThrThr AlaIlePro LeuPheAlaVal GlnAsnTyr GlnValPro LeuLeuSerVal

 481 TACGTGCAGG CCGCCAACCT GCACCTGAGC GTGCTGCGCG ACGTCAGCGT GTTCGGCCAG
     TyrValGln AlaAlaAsn LeuHisLeuSer ValLeuArg AspValSer ValPheGlyGln

 541 CGCTGGGGCT TCGACGCCGC CACCATCAAC AGCCGCTACA ACGACCTGAC CCGCCTGATC
     ArgTrpGly PheAspAla AlaThrIleAsn SerArgTyr AsnAspLeu ThrArgLeuIle

 601 GGCAACTACA CCGACCACGC CGTGCGCTGG TACAACACCG GCCTGGAGCG CGTGTGGGGT
     GlyAsnTyr ThrAspHis AlaValArgTrp TyrAsnThr GlyLeuGlu ArgValTrpGly

 661 CCCGACAGCC GCGACTGGAT CAGGTACAAC CAGTTCCGCC GCGAGCTGAC CCTGACCGTG
     ProAspSer ArgAspTrp IleArgTyrAsn GlnPheArg ArgGluLeu ThrLeuThrVal

 721 CTGGACATCG TGAGCCTGTT CCCCAACTAC GACAGCCGCA CCTACCCCAT CCGCACCGTG
     LeuAspIle ValSerLeu PheProAsnTyr AspSerArg ThrTyrPro IleArgThrVal

 781 AGCCAGCTGA CCCGCGAGAT TTACACCAAC CCCGTGCTGG AGAACTTCGA CGGCAGCTTC
     SerGlnLeu ThrArgGlu IleTyrThrAsn ProValLeu GluAsnPhe AspGlySerPhe

 841 CGCGGCAGCG CCCAGGGCAT CGAGGGCAGC ATCCGCAGCC CCCACCTGAT GGACATCCTG
     ArgGlySer AlaGlnGly IleGluGlySer IleArgSer ProHisLeu MetAspIleLeu

 901 AACAGCATCA CCATCTACAC CGACGCCCAC CGCGGCGAGT ACTACTGGAG CGGCCACCAG
     AsnSerIle ThrIleTyr ThrAspAlaHis ArgGlyGlu TyrTyrTrp SerGlyHisGln

 961 ATCATGGCCA GCCCCGTCGG CTTCAGCGGC CCCGAGTTCA CCTTCCCCCT GTACGGCACC
     IleMetAla SerProVal GlyPheSerGly ProGluPhe ThrPhePro LeuTyrGlyThr

1021 ATGGGCAACG CTGCACCTCA GCAGCGCATC GTGGCACAGC TGGGCCAGGG AGTGTACCGC
     MetGlyAsn AlaAlaPro GlnGlnArgIle ValAlaGln LeuGlyGln GlyValTyrArg

1081 ACCCTGAGCA GCACCCTGTA CCGTCGACCT TTCAACATCG GCATCAACAA CCAGCAGCTG
     ThrLeuSer SerThrLeu TyrArgArgPro PheAsnIle GlyIleAsn AsnGlnGlnLeu

1141 AGCGTGCTGG ACGGCACCGA GTTCGCCTAC GGCACCAGCA GCAACCTGCC CAGCGCCGTG
     SerValLeu AspGlyThr GluPheAlaTyr GlyThrSer SerAsnLeu ProSerAlaVal

1201 TACCGCAAGA GCGGCACCGT GGACAGCCTG GACGAGATCC CCCCTCAGAA CAACAACGTG
     TyrArgLys SerGlyThr ValAspSerLeu AspGluIle ProProGln AsnAsnAsnVal
```

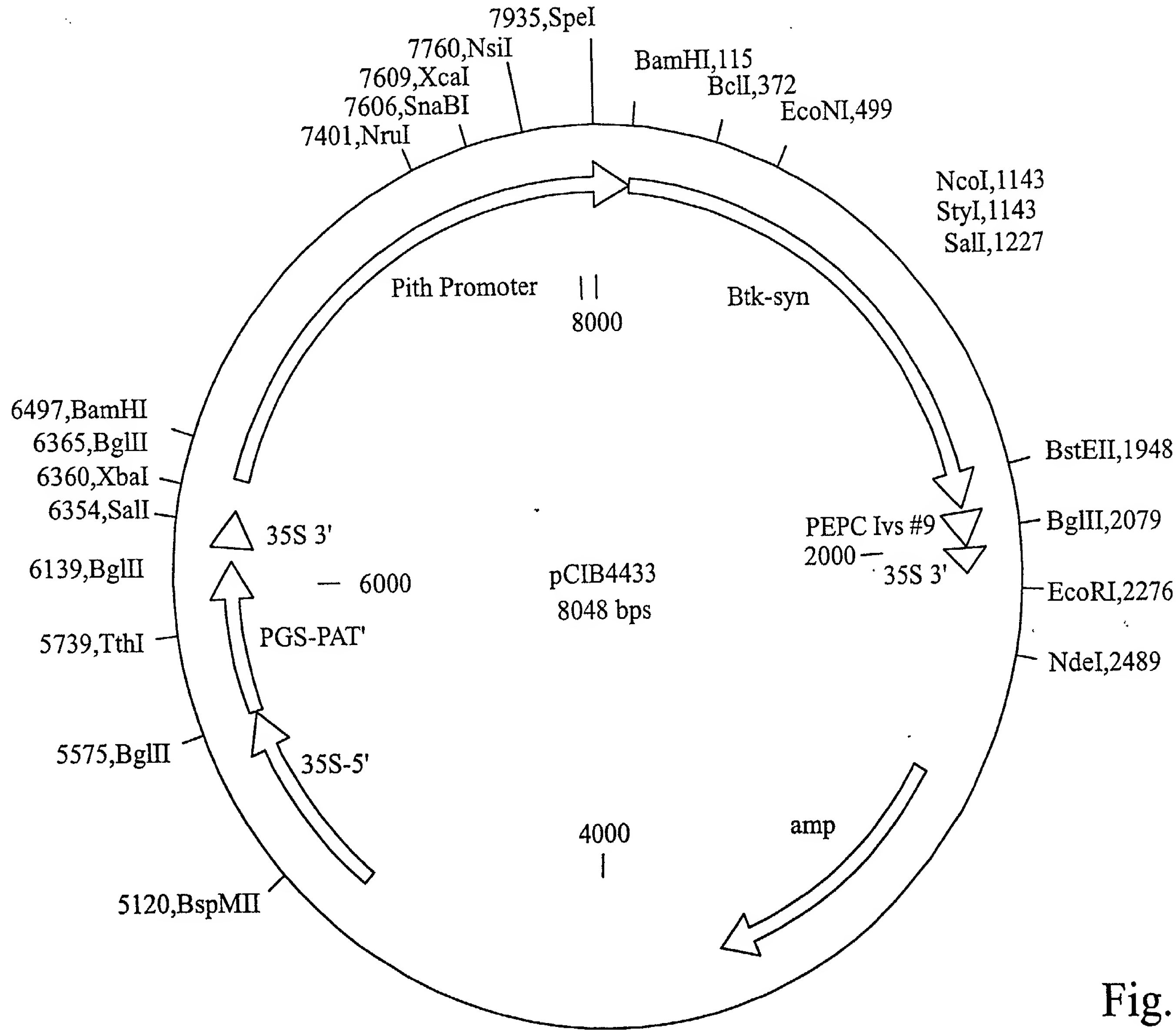

7935,SpeI
7760,NsiI
7609,XcaI
7606,SnaBI
7401,NruI
BamHI,115
BclI,372
EcoNI,499
NcoI,1143
StyI,1143
SalI,1227
Pith Promoter
8000
Btk-syn
6497,BamHI
6365,BglII
6360,XbaI
6354,SalI
6139,BglII
5739,TthI
5575,BglII
5120,BspMII
35S 3'
6000
PGS-PAT'
35S-5'
pCIB4433
8048 bps
4000
amp
PEPC Ivs #9
2000
35S 3'
BstEII,1948
BglII,2079
EcoRI,2276
NdeI,2489

Fig. 36

Fig. 35E

pol CDPK gene Map (1 > 4165) Site and Sequence

TCAGCCGACAAACTAAACTATAGAAACCACATCATGATATCAAATTTTGAGGTGGCGGTGCTACAGAAATAGAACCCAGT 3600

INTRON 6

ACACCAAAATGACTAACTTGTCATGATTAGTTGTTCCTCGTAACTGAACATTTGTGTTCTTAGTTTCTTATTGTTAAACC 3680

INTRON 6

AAAGACTTAAATTCACTTTTGCACATGCAGGATGGAAGGATAGATTATTCAGAGTTTGTGGCGATGATGAGGAAAGGGAC 3760

INTRON 6 EXON 7

GGCTGGTGCCGAGCCAATGAACATCAAGAAGAGGCGAGACATAGTCCTATAGTGAAGTGAAGCAGWAAGTGTGTAATGTA 3840

EXON 7

ATGTGTATAGCAGCTCAAACAAGCAAATTTGTACATCTGTACACAAATGCAATGGGGTTACTTTTGCAACTTAGTTCATG 3920

GATGGTTGTGTACGTTGTGCTATTGATTGCAAGTGATTTGAAAGACATGCATACTTAGGAACTGAGAAAGATAGATCTAC 4000

TACTGCTAGAGACAGAACAATAGGATKKYAATTCAGYAAGTGYGTATTTCAGAAGACTACAGCTGGCATCTATTATTCTC 4080

ATTGTCCTCGCAAAAATACTGATGATGCATTTGAGAGAACAATATGCAACAAGATCGAGCTCCCTATAGTGAGTCGTATT 4160

AGGCC

4165

Fig. 35D

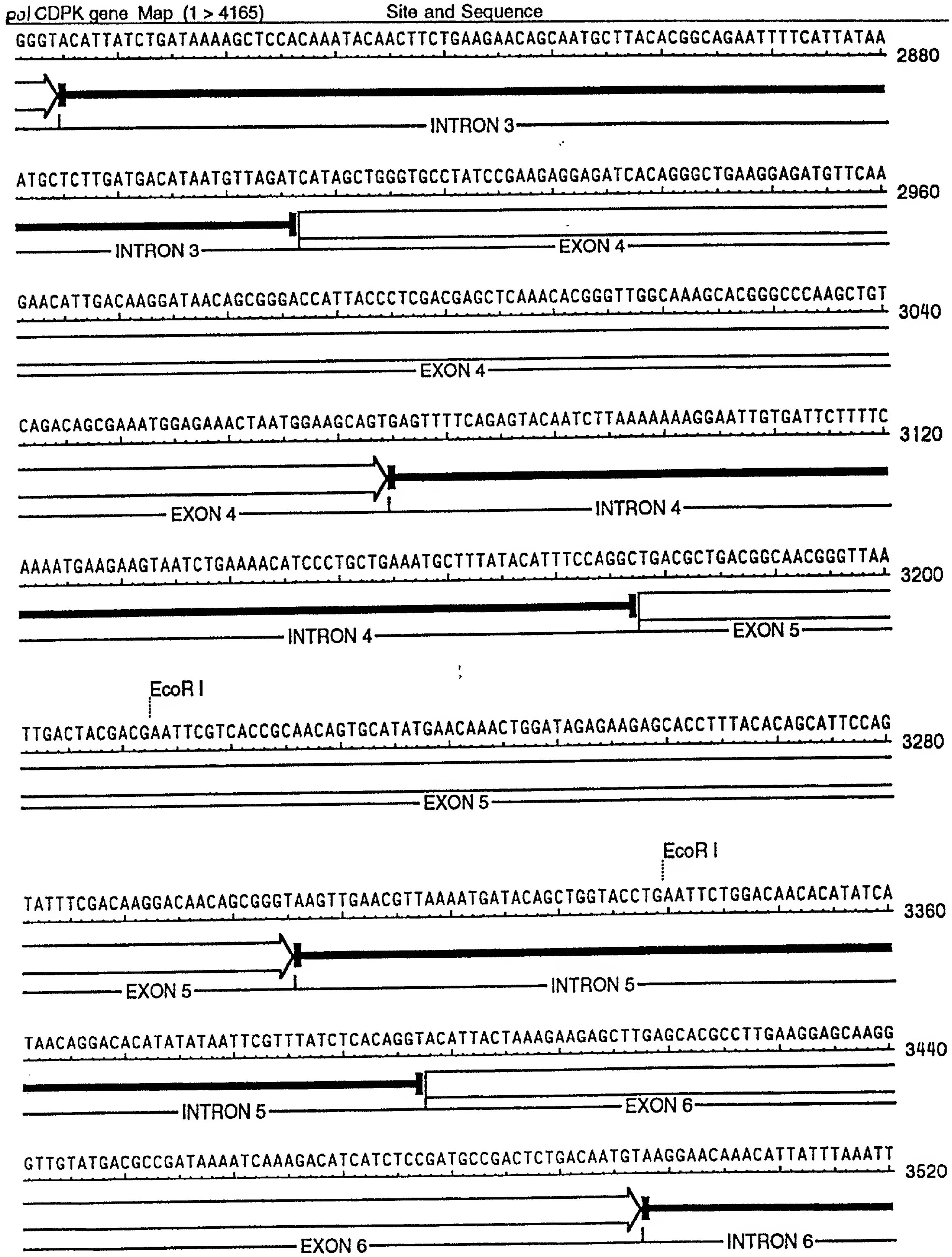

pol CDPK gene Map (1 > 4165)
Site and Sequence
GGGTACATTATCTGATAAAAGCTCCACAAATACAACTTCTGAAGAACAGCAATGCTTACACGGCAGAATTTTCATTATAA 2880
INTRON 3
ATGCTCTTGATGACATAATGTTAGATCATAGCTGGGTGCCTATCCGAAGAGGAGATCACAGGGCTGAAGGAGATGTTCAA 2960
INTRON 3
EXON 4
GAACATTGACAAGGATAACAGCGGGACCATTACCCTCGACGAGCTCAAACACGGGTTGGCAAAGCACGGGCCCAAGCTGT 3040
EXON 4
CAGACAGCGAAATGGAGAAACTAATGGAAGCAGTGAGTTTTCAGAGTACAATCTTAAAAAAAGGAATTGTGATTCTTTTC 3120
EXON 4
INTRON 4
AAAATGAAGAAGTAATCTGAAAACATCCCTGCTGAAATGCTTTATACATTTCCAGGCTGACGCTGACGGCAACGGGTTAA 3200
INTRON 4
EXON 5
EcoR I
TTGACTACGACGAATTCGTCACCGCAACAGTGCATATGAACAAACTGGATAGAGAAGAGCACCTTTACACAGCATTCCAG 3280
EXON 5
EcoR I
TATTTCGACAAGGACAACAGCGGGTAAGTTGAACGTTAAAATGATACAGCTGGTACCTGAATTCTGGACAACACATATCA 3360
EXON 5
INTRON 5
TAACAGGACACATATATAATTCGTTTATCTCACAGGTACATTACTAAAGAAGAGCTTGAGCACGCCTTGAAGGAGCAAGG 3440
INTRON 5
EXON 6
GTTGTATGACGCCGATAAAATCAAAGACATCATCTCCGATGCCGACTCTGACAATGTAAGGAACAAACATTATTTAAATT 3520
EXON 6
INTRON 6

Fig. 35C

pol CDPK gene Map (1 > 4165) Site and Sequence

Ava I

CGTGCAGATCGTGCACACCTGCCACTCCATGGGGGTGATGCACCGGGACATCAAGCCCGAGAACTTCCTGCTGCTCAGCA 2160

EXON 1

AGGACGAGGACGCGCCGCTCAAGGCCACCGACTTCGGCCTCTCCGTCTTCTTCAAGGAGGGCGAGCTGCTCAGGGACATC 2240

EXON 1

Ava I

GTCGGCAGCGCCTACTACATCGCGCCCGAGGTGCTCAAGAGGAAGTACGGCCCGGAGGCCGACATCTGGAGCGTCGGCGT 2320

EXON 1

Bam H I

CATGCTCTACATCTTCCTCGCCGGCGTGCCTCCCTTCTGGGCAGGTCGGATCCGTCCGTGTTCGTCCTAGACGATATACA 2400

EXON 1 INTRON 1

GAACCCGACGATGGATTTGCTTCTCAGCCCTGTTCTTGCATCACCAGAGAACGAGAACGGCATCTTCACCGCCATCCTGC 2480

INTRON 1 EXON 2

GAGGGCAGCTTGACCTCTCCAGCGAGCCATGGCCACACATCTCGCCGGGAGCCAAGGATCTCGTCAAGAAGATGCTCAAC 2560

EXON 2

ATCAACCCCAAGGAGCGGCTCACGGCGTTCCAGGTCCTCAGTAAGTACCCAGATCGTTGCTGTCATACACTCATATGAAT 2640

EXON 2 INTRON 2

TGTATCGTTCATGAGCAACGATCGAGCGGATTTGGTGAACTTGTAGATCACCCATGGATCAAAGAAGACGGAGACGCGCC 2720

INTRON 2 EXON 3

TGACACGCCGCTTGACAACGTTGTTCTCGACAGGCTCAAGCAGTTCAGGGCCATGAACCAGTTCAAGAAAGCAGCATTGA 2800

EXON 3

Fig. 35B

pol CDPK gene Map (1 > 4165) Site and Sequence

Xba I

CAAAATCTAGAAACCTTTTTTTTTCCTCCCGATACGCCCCTCCATCTCTCGCCGTTCATGTCCGTGGCTGGCTGCCCTCC 1440

mRNA start

GTGGGAGCAGGCGGCCGCACTCGTTCCCCGCCGCAGCCATGGGCCAGTGCTGCTCCAAGGGCGCCGGAGAGGCCCCGCCA 1520

EXON 1

CCGAGGCGCCAAACGGCAGGCGCCAAGCCGCGGGCGTCCGCGAACAACGCCGACGGACAACGGGCGTCGTCCTCGTCCGC 1600

EXON 1

GGTGGCTGCTGCCGCTGCTGCTGCCGGTGGTGGTGGCGGCGGCACGACGAAGCCGGCCTCACCCACCGGCGGCGCCAGGG 1680

EXON 1

CCAGCTCCGGCAGCAAACCGGCGGCGGCCGTGGGCACGGTGCTGGGCCGGCCCATGGAGGACGTGCGCGCGACCTACTCG 1760

EXON 1

Ava I

ATGGGCAAGGAGCTCGGGCGCGGGCAGTTCGGCGTGACGCACCTGTGCACGCACCGGACGAGCGGCGAGAAGCTGGCGTG 1840

EXON 1

CAAGACGATCGCGAAGCGGAAGCTGGCGGCCAGGGAGGACGTGGACGACGTGCGGCGGGAGGTGCAGATCATGCACCACC 1920

EXON 1

TCTCCGGCCAGCCCAACGTGGTGGGCCTCCGCGGCGCGTACGAGGACAAGCAGAGCGTGCACCTCGTCATGGAGCTGTGC 2000

EXON 1

Ava I

GCGGGCGGGGAGCTCTTCGACCGCATCATCGCCCGGGGCCAGTACACGGAGCGCGGCGCCGCGGAGCTGCTGCGCGCCAT 2080

EXON 1

Fig. 35A

pol CDPK gene Map (1 > 4165)   Site and Sequence
Enzymes :   6 of 198 enzymes (Filtered)
Settings :   Circular, Certain Sites Only, Standard Genetic Code

TTAGTAACACCTCTCCAATCGCTTGGGTTGGCACATTCTTAGCTTTTATCACATTTTAAGAAATAGAGTTCACCACCTTC 80

AAAATATGCCTATACAATGAATGATGCTTGGATGCAATATAGCTAGATTCAACTAGCTATATATGGTCAATAGAACCCTG 160

TGAGCACCTCACAAACACGACTTCAATTTTGAGACCCTAAGCGAGTAAATGGTTAAAGTCCTCTTATTATTAGTCTTAGG 240

ACTTCTCCTTGCTAAATGCTTGTCAGCGATCTATATATCTTCCCCACTGCGGGAGATACTATATATAGGGCCTTGGACCT 320

CTAGGGTATCTCAAAGGCCTAGTCACAACAATTCTCAACAGTATTTAATTTTATACATGTATGAACAGTGTAGGAATTTG 400

AGTGCCCAACCCAAGAGTGGGAGGTGTAAATTGGGTAGCTAAACTTAAATAGGGCTCTTCTTATTTAGGTTTATCTAGTC 480

TCTACTTAGACTAATTCAGAAAGAATTTTACAACCTATGGTTAATCATATCTCTAGTCTAAGCAAATTTAGGAAAGTTAA 560

AAGCACACAATTAGGCACATGTGAAAGATGTGTATGGTAAGTAAAAGACTTATAAGGAAAAAGTGGGTGAATCCTCAAGA 640

TGTGGTGGTATATCCCAATGATATTAGATGCCAGAATATAGGGGGGAAATCGATGTATACCATCTCTACCAGGATACCTG 720

TGCGGACTGTGCAACTGACACATGGACCATGGTGTCTTCTTAGATTTGGTTATTAGCTAATTGCGCTACAACTTGTTCAA 800

GGCTAGACCAAATTAAAAAACTAATATTAAACATAAAAAGTTAGGCAAACTATAGTAAATTATGCAGCGATCCAACAACA 880

AGCCATGTCTCGTGGGTCATGAGCCACGCGTCGGCCATACACCCACATGATGTTTCCATACGGATGGTCCTTATGCAATT 960

TTGTCTGCAAAACACAAGCCTTAATACAGCCACGCGACAATCATGGAAGTGGTCGTTTTAGGTCCTCATCATGAAGTTCA 1040

GGGAAAACGCATCAAATGTAATGCAGAGAAATGGTATTTCTTCTCTTGTAAATCAGGGAGAGGAGTACCATCAGTACAGA 1120

EcoR I

TTCAGAATCAGAATTCAGTCTTCCAACGACAATAATCGCAGCATCTTGTAAAAATTTGCAGAAACTTCTGTTTGACTTGT 1200

AGCCCTGACCTTTGCAAATATTTGAAGTTGTGCCTGCTGACACAACTTCAATCTGGAAGTGCTGTTGATCAGTTTTGCCA 1280

GAAACAGCAAGCAGCCTATATATATCTGTCACGAGACACCCTGCCGCCCTCTTCTTTCCCGCCATTCCCTCCCTACCCTT 1360

Fig. 34

Lipman-Pearson Protein Alignment
Gap Penalty: 2; Gap Length Penalty: 12

Seq1 pol CDPK ptn	Seq2 soybean CDPK ptn	Similarity Index	Gap Number	Gap Length	Consensus Length
1>551	1>509	62.4	1	1	464

```
pol CDPK ptn       VLGRPMEDVRATYSMGKELGRGQFGVTHLCTHRTSGEKLACKTIAKRKLAAREDVDDVRREVQIMHHLSG 150
                   || :. :::|..|.:|:.||:||||.|  ||:|:||.|:|||:|:||||  :|| :||:||:|||||||.
soybean CDPK ptn   VLPQRTQNIREVYEVGRKLGQGQFGTTFECTRRASGGKFACKSIPKRKLLCKEDYEDVWREIQIMHHLSE 91

pol CDPK ptn       QPNVVGLRGAYEDKQSVHLVMELCAGGELFDRIIARGQYTERGAAELLRAIVQIVHTCHSMGVMHRDIKP 220
                   ::||| : |:|||. :||||||||.||||||||:.:|:|:|| || |:::||::|::|||:||||||:||
soybean CDPK ptn   HANVVRIEGTYEDSTAVHLVMELCEGGELFDRIVQKGHYSERQAARLIKTIVEVVEACHSLGVMHRDLKP 161

pol CDPK ptn       ENFLLLSKDEDAPLKATDFGLSVFFKEGELLRDIVGSAYYIAPEVLKRKYGPEADIWSVGVMLYIFLAGV 290
                   ||||: : |||| |||||||||||:| || : |:|||:||:|||||:: ||||:|:||.||:|||:|:||
soybean CDPK ptn   ENFLFDTIDEDAKLKATDFGLSVFYKPGESFCDVVGSPYYVAPEVLRKLYGPESDVWSAGVILYILLSGV 231

pol CDPK ptn       PPFWAENENGIFTAILRGQLDLSSEPWPHISPGAKDLVKKMLNINPKERLTAFQVLNHPWIKEDGDAPDT 360
                   ||||||:| ||| .|| |:||: ||||| || :||||::|||: |||.|||| :||.|||| :|. |||.
soybean CDPK       PPFWAESEPGIFRQILLGKLDFHSEPWPSISDSAKDLIRKMLDQNPKTRLTAHEVLRHPWIVDDNIAPDK 301

pol CDPK ptn       PLDNVVLDRLKQFRAMNQFKKAALRIIAGCLSEEEITGLKEMFKNIDKDNSGTITLDELKHGLAKHGPKL 430
                   |||:.||.|||||.|||::|| |||:||. ||||||.||||:|| ||.|||||||:||||:|| : |:.|
soybean CDPK ptn   PLDSAVLSRLKQFSAMNKLKKMALRVIAERLSEEEIGGLKELFKMIDTDNSGTITFDELKDGLKRVGSEL 371

pol CDPK ptn       SDSEMEKLMEAADADGNGLIDYDEFVTATVHMNKLDREEHLYTAFQYFDKDNSGYITKEELEHALKEQGL 500
                    :||:..||:||| | :| |||:||::||||:|||:|||:| :|| |||||.||||| :|:::| |: ||
soybean CDPK ptn   MESEIKDLMDAADIDKSGTIDYGEFIAATVHLNKLEREENLVSAFSYFDKDGSGYITLDEIQQACKDFGL 441

pol CDPK ptn       YDADKIKDIISDADSDNDGRIDYSEFVAMMRKGTAGAEPMNIKK   544
                    |. .|.|:|.: | ||||:|||:||.||||||.:| ....::|
soybean CDPK ptn   -DDIHIDDMIKEIDQDNDGQIDYGEFAAMMRKGNGGIGRRTMRK   484
```

Fig. 33

Lipman-Pearson Protein Alignment
Gap Penalty: 2; Gap Length Penalty: 12

Seq1 pol CDPK ptn	Seq2 humcama ptn	Similarity Index	Gap Number	Gap Length	Consensus Length
1>551	1>150	40.3	2	2	142

```
pol CDPK ptn    LSEEEITGLKEMFKNIDKDNSGTITLDELKHGLAKHGPKLSDSEMEKLMEAADADGNGLIDYDEFVTATV 460
                L:EE:I:.:KE F. :DKD..GTIT .EL   : . G.: :::E::.:::..DADGNG ID: EF:T  .
humcama  ptn    LTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSLGQNPTEAELQDMINEVDADGNGTIDFPEFLTMMA 74


pol CDPK ptn    HMNKL-DREEHLYTAFQYFDKDNSGYITKEELEHALKEQGLYDADKIKDI-ISDADSDNDGRIDYSEFVA 528
                : M:. D.EE:: .AF: FDKD.:GYI: .EL H.:.: G   ..:.:.::I.:AD D.DG:::Y.EFV.
humcama  ptn    RKMKDTDSEEEIREAFRVKDKDGNGYISAAELRHVMTNLGEKLTDEEVDEMIREADIDGDGQVNYEEFVQ 144


pol CDPK ptn    MM 530
                MM
humcama  ptn    MM 146
```

Fig. 32

```
Lipman-Pearson Protein Alignment
Gap Penalty: 2;  Gap Length Penalty: 12
Seq1             Seq2           Similarity    Gap     Gap  Consensus
pol CDPK ptn     rat pk2 ptn         Index Number  Length     Length
1>551            1>528                36.5      4       4        297
```

```
pol CDPK ptn  YSMGKELGRGQFGVTHLCTHRTSGEKLACKTIAKRKLAAREDVDDVRREVQIMHHLSGQPNVVGLRGAYE 162
              Y : .ELG:G.F:V.: C..:TS.:. A K.I..:KL:AR:: :.: RE.:I : L. :PN:V L::: .
rat pk2  ptn  YQLFEELGKGAFSVVRRCVKKTSTQEYAAKIINTKKLSARDH-QKLEREARICRLLK-HPNIVRLHDSIS 81

pol CDPK ptn  DKQSVHLVMELCAGGELFDRIIARGQYTERGAAELLRAIVQIVHTCHSMGVMHRDIKPENFLLLSKDEDA 232
              :.   .LV.:. :GGELF: I:AR. Y:E :A:: ::.I:: V:  H  :::HRD:KPEN:LL SK .:A
rat pk2  ptn  EEGFHYLVFDLVTGGELFEDIVAREYYSEADASHCIHQILESVNHIHQHDIVHRDLKPENLLLASKCKGA 151

pol CDPK ptn  PLKATDFGLSVFFK-EGELLRDIVGSAYYIAPEVL-KRKYGPEADIWSVGVMLYIFLAGVPPFWAENENG 300
              ::K :DFGL::  :..:  : ::.G:: Y::PEVL::. YG  .DIW: GV:LYI:L.G PPFW.E:::
rat pk2  ptn  AVKLADFGLAIEVQGEQQAWFGFAGTPGYLSPEVLRKDPYGKPVDIWACGVILYILLVGYPPFWDEDQHK 221

pol CDPK ptn  IFTAILRGQLDLSSEPWPHISPGAKDLVKKMLNINPKERLTAFQVLNHPWIKEDGDAPDTPLDNVVLDRL 370
              :: .I  G. D::S  W  ::P.AK:L:::ML.INP .R:TA Q.L:HPW: : :..:.    : .:: L
rat pk2  ptn  LYQQIKAGAYDFPSPEWDTVTPEAKNLINQMLTINPAKRITADQALKHPWVCQRSTVASMMHRQETVECL 291

pol CDPK ptn  KQFRAMNQFKKAALRII 387
              ::F.A..::K A L .:
rat pk2  ptn  RKFNARRKLKGAILTTM 308
```